University of Alberta

3-D LEAST-SQUARES WAVE-EQUATION AVP / AVA MIGRATION OF COMMON-AZIMUTH DATA

by

Juefu Wang ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

in

Geophysics

Department of Physics

Edmonton, Alberta
Fall 2005

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

0-494-08749-8

Your file Votre référence
ISBN:
Our file Notre reterence
ISBN:

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

To
Minzi

Abstract

This thesis presents a 3-D least-squares wave-equation migration that yields regularized common image gathers for amplitude versus angle (AVA) analysis. It is an extension of 2-D least-squares migration to the 3-D case using the common-azimuth approximation.

The thesis generalizes least-squares wave-equation migration with various types of regularization. In particular, an efficient preconditioning strategy is adopted to decrease the cost of the iterative inversion. The thesis also proposes a new scheme that combines a smoothness regularization in the ray parameter direction with a sparseness regularization in the depth direction to further improve the resolution of seismic images. The problem is solved by an iterative re-weighted least-squares conjugate gradients algorithm (IRLS).

Extensive tests on synthetic and real data show that regularized iterative migration/ inversion enhances event continuity in common image gathers. More importantly, the method is validated by a careful comparison of the inverted common image gathers and synthetics obtained from well log data. In addition, the vertical resolution of the inverted image is also improved as a consequence of increased coherence in common image gathers, sparse regularization in the depth direction, and by implicitly introducing migration deconvolution in the inversion.

Acknowledgements

I would like to thank my supervisor, Dr. Mauricio D. Sacchi for his guidance to the world of inversion. The inverse theory changed my way of thinking and opened my mind to a new charming world full of inverse problems. I should also thank Dr. Henning Kuehl for sharing his experience of least-squares migration. Thanks to my SAIG colleagues for their friendship, intellectual support and valuable suggestions in the past few years. Thanks to my fellow Geophysics soccer maniacs for the fun games. Finally I should thank my family for their support and always having confidence in me.

Contents

1 Introduction 1

2 Seismic modelling and migration/the adjoint 8

2.1 Introduction 8

2.2 Kirchhoff modelling and migration/the adjoint 9

2.2.1 WKBJ approximation 9

2.2.2 Green's function for linear partial differential operators and derivation of Kirchhoff modelling 11

2.2.3 Migration/the adjoint and a preview of inversion 15

2.3 Some comments on the Kirchhoff operator 16

2.4 Wave-equation modelling/migration 22

2.4.1 Two-way versus one-way wave-equation representation 22

2.4.2 Two-way vs. one-way wave-equation modelling/migration 27

2.5 Angle-dependent reflectivity 33

2.6 Angle-domain imaging condition 40

2.7 DSR wave-equation modelling/migration 45

2.7.1 Theory of DSR modelling/migration 45

2.7.2 A toy example of AVP migration 52

2.8 Common-azimuth approximation 55

2.9 Implementation for laterally variant media 61

2.10 Summary 68

3 Regularized common-azimuth least-squares AVP migration 70

3.1 Introduction 70

3.2 Regularized least-squares migration 71

3.3 Conjugate gradients algorithm 74
3.4 Least-squares migration with smooth regularization 76
3.5 Preconditioned implementation of RLSM 77
3.6 A toy example: inversion of a smooth model 78
3.7 Sparse least-squares migration 85
3.8 Hyperparameter tests 87
3.9 A simple example: multi-channel deconvolution 89
3.10 True amplitude migration 91
3.11 Summary 94

4 Synthetic data examples **95**
4.1 Migration 95
4.2 Least-squares migration 111
  4.2.1 Least-square migration with smooth regularization 111
  4.2.2 Sparse least-squares migration 124
4.3 Summary 143

5 Field data example **144**
5.1 Introduction 144
5.2 Erskine data from WCSB 145
5.3 Migration vs. LSM with smooth regularization 146
5.4 PLSM vs. SLSM 156
5.5 Stacked cubes and depth slices 160
5.6 Summary 160

Discussion and conclusions **163**

References **167**

A Propagation matrix for two-way wave equation **174**

B Relationship between pressure, acoustic impedance and particle velocity for plane waves **176**

C Imaging Jacobian for common-azimuth migration **178**

List of Tables

2.1 Some migration operators and their wavenumber values 51
2.2 Model parameters for the 2-D synthetic data. 53

4.1 Model parameters for the 3-D synthetic data. 112

List of Figures

1.1 Incident P wave and the reflected and transmitted waves at an interface between two media. 2
1.2 A simple example of NMO correction. 4

2.1 Diagram for deriving the Kirchhoff integral operator. 12
2.2 An example of 1-D Kirchhoff inversion. 17
2.3 Velocity profile A-A′ from the SEG/EAGE data. 18
2.4 Zero-offset data for the profile A-A′ from SEG/EAGE model. 19
2.5 Migrated images by Kirchhoff method using different macro velocity models. 21
2.6 The scheme for one-way wave-equation modelling. 30
2.7 Migration of a single shot gather 33
2.8 Diagram to derive angle-dependent reflectivity of a plane wave. 34
2.9 A common midpoint gather, the inverted radon panels, reconstructed data and the residuals. 41
2.10 Derivation of apparent velocities. 46
2.11 Schemes of Double-Square-Root (DSR) wave equation modeling and migration. 51
2.12 Ray parameter dependent image gathers by wave-equation migration. . . 53
2.13 Ray parameter dependent image gathers by wave-equation migration. . . 54
2.14 AVA curves by different imaging methods. 56
2.15 The scheme of the split-step Fourier migration. 64
2.16 The scheme of the extended split-step Fourier migration (split-step PSPI). 65
2.17 The scheme of the PSPI migration. 66

2.18 Velocity distribution of the 2-D SEG/EAGE salt model at depth $z = 2520ft$. 68

3.1 A smooth model and the simulated noisy data. 79
3.2 Reconstructed model (Blue) vs. the true model (Red) at different regularized CG iterations. 81
3.3 Reconstructed model (Blue) vs. the true model (Red) at different preconditioned CG iterations. 82
3.4 Data misfit comparison between the regularized CG and the preconditioned CG 83
3.5 Eigenvalue comparison between the regularized CG and the preconditioned CG 83
3.6 The reconstructed model vs. the true model. 84
3.7 Sparse reflectivity model and the simulated noisy seismic data. 88
3.8 True reflectivity model vs. inverted models by using a constant trade-off parameter. 89
3.9 True reflectivity model vs. inverted models by using a constant scaling parameter. 90
3.10 True reflectivity model vs. inverted models by different sets of hyperparameters with a constant product. 90
3.11 A multi-channel deconvolution example to compare linear inversion and non-linear inversion. 92

4.1 3-D SEG/EAGE salt model. 96
4.2 Migrated image of the 'exploding-reflector' data by split-step zero-offset migration. 98
4.3 Migrated image of the 'exploding-reflector' data by split-step PSPI zero-offset migration. 98
4.4 Migrated image of the 'exploding-reflector' data by PSPI zero-offset migration. 99
4.5 Migrated image of the 'exploding-reflector' data by Fourier Finite Difference zero-offset migration. 99

4.6 The geometry of marine data acquisition for 3-D SEG/EAGE salt model. . 100
4.7 The survey area of 3-D SEG/EAGE Narrow-azimuth data for the salt model. 101
4.8 Image comparison between 2-D and 3-D migration at $x = 5460$ m 104
4.9 Stacked images at $x = 6760$ m, obtained with 3-D common-azimuth migration and 3-D shot migration. 105
4.10 Stacked images at $x = 8440$ m, obtained with 3-D common-azimuth migration and 3-D shot migration. 106
4.11 Stacked images at $y = 3380$ m, obtained with 3-D common-azimuth migration and 3-D shot migration. 107
4.12 The velocity and density profiles of the Marmousi model. 109
4.13 Migrated images of the Marmousi data. 110
4.14 Common image gathers of the Marmousi data at $x = 7500$ m 111
4.15 Comparison between observed and reconstructed data. 113
4.16 CIGs with varying CG iterations. 115
4.17 Extracted AVA curves. 116
4.18 CIGs by Regularized Least-Squares Migration (RLSM) and Preconditioned Least-Squares Migration (PLSM). 117
4.19 Extracted AVA curves. 118
4.20 Data misfit of two inversion methods. 118
4.21 CIGs by migration and Preconditioned Least-Squares Migration (PLSM). 121
4.22 Extracted AVA curves of the noisy incomplete data. 122
4.23 Data misfit of the preconditioned least-squares migration applied to the noisy incomplete data. 122
4.24 Comparison of the observed, reconstructed data and the residual. 123
4.25 Comparison of the predicted noise, the real additive noise and the residual. 123
4.26 The geological structure of the wedge model. 124
4.27 Synthesized common midpoint gathers at midpoint $x = 500$ m of the wedge model. 125
4.28 Stacked images of the wedge model, inverted from noise-free dataset 1. . 126
4.29 Common image gathers (CIGs) and AVA curves at $x = 500$ m for the wedge model. 127

4.30 Data misfit of two inversion methods. 128

4.31 The original data, reconstructed data by PLSM and the residual at $x = 500$ m. 129

4.32 The original data, reconstructed data by SLSM and the residual at $x = 500$ m. 129

4.33 Stacked images of the wedge model, inverted from noise-free dataset 2. . 130

4.34 Common image gathers (CIGs) and AVA curves of the noise-free dataset 2 at $x = 500$ m of the wedge model. 133

4.35 Common image gathers (CIGs) and AVA curves at $x = 500$ m of the wedge model. 134

4.36 Common image gathers (CIGs) and AVA curves of the noisy dataset 1 at $x = 500$ m of the wedge model. 135

4.37 Comparison between observed, reconstructed data and the residual. . . . 136

4.38 Comparison between observed, reconstructed data and the residual. . . . 136

4.39 Common image gathers of migration, the Preconditioned Least-Squares Migration (PLSM) and the Sparse Least-Squares Migration (SLSM) for the incomplete Marmousi data. 138

4.40 Zoomed-in common image gathers of the Marmousi data. 138

4.41 Stacked image by migration of the incomplete Marmousi data. 139

4.42 Stacked image by the preconditioned least-squares migration. 140

4.43 Stacked image by the sparse least-squares migration. 140

4.44 Image of constant ray parameter ($p_{h_0} = 0\mu s/m$) by migration of the incomplete Marmousi data. 141

4.45 Image of constant ray parameter ($p_{h_0} = 0\mu s/m$) by the preconditioned least-squares migration of the incomplete Marmousi data. 142

4.46 Image of constant ray parameter ($p_{h_0} = 0\mu s/m$) by the sparse least-squares migration of the incomplete Marmousi data. 142

5.1 Offset distribution per CMP bin for the Erskine dataset (WCSB). 145

5.2 Comparison between observed and reconstructed data at inline No.10. . . 147

5.3 Common image gathers at crossline No. 10, inline No. 71. 148

5.4 Normalized residual norm $||\mathbf{W}(\mathbf{Lm} - \mathbf{d})||^2$ versus CG iteration. 149
5.5 Common image gathers at crossline No. 10, inline No. 71. 150
5.6 Detailed image of inline No.71 with CIG at crossline No. 10. 151
5.7 Stacked images of crossline No. 24. 152
5.8 Depth correlation of log traces. 154
5.9 Strata correlation for the survey area of the Erskine dataset. 155
5.10 CIG at inline No. 76, crossline No. 24. 156
5.11 Comparison between the synthetic CIG and the inverted CIG. 157
5.12 Comparison between the synthetic AVA and the inverted AVA. 157
5.13 Stacked images obtained by the preconditioned least-squares migration (PLSM) and the sparse least-squares migration (SLSM). 158
5.14 Inverted common image gathers by the preconditioned least-squares migration (PLSM) and the sparse least-squares migration (SLSM). 159
5.15 Stacked cubes obtained by migration, regularized least-squares migration (RLSM), preconditioned least-squares migration (PLSM) and sparse least-squares migration (SLSM). 161
5.16 Stacked depth slices at $z = 1800$ m, obtained by migration, regularized least-squares migration (RLSM), preconditioned least-squares migration (PLSM) and sparse least-squares migration (SLSM). 162

List of symbols and abbreviations

Numerics and rock physics

u	displacement field
∇	Laplacian operator
ω	temporal frequency
k	wave number
p	ray parameter
k_{m_x}	inline midpoint wavenumber
k_{m_y}	crossline midpoint wavenumber
k_{h_x}	inline offset wavenumber
k_{h_y}	crossline offset wavenumber
p_{h_x}	inline offset ray parameter
p_{h_y}	crossline offset ray parameter
τ	traveltime
A_j	j th order amplitude
G	Green's function
$\mathbf{x}_s$	source vector
$\mathbf{x}_g$	receiver vector
$\mathbf{x}$	image point or midpoint vector
R	reflectivity or regularization function
θ	incident angle or structural dip angle
γ	incident angle
ϕ	azimuth angle
c	wave velocity
ρ	density
K	bulk compression modulus
$\mathbf{L}$	modelling or wave composition operator
$\mathbf{L}^{-1}$	wave decomposition operator
$\mathbf{m}$	the earth model in terms of common image gathers
$\mathbf{d}$	seismic data
$\mathbf{n}$	additive noise
$\mathbf{W}$	data sampling matrix
$\mathbf{W}_s$	source signature (wavelet)
$\mathbf{D}_{1h_x}$	first order derivative operator along the inline offset ray parameter direction
$\mathbf{L}_\omega$	bandlimited wave propagator
λ	trade-offset parameter for inversion
δ	delta function or scaling factor for sparse inversion
P^+	downgoing wave
P^-	upgoing wave

Abbreviations

AVA	Amplitude Versus Angle
AVAZ	Amplitude Versus Angle and aZimuth
AVO	Amplitude Versus Offset
AVP	Amplitude Versus ray Parameter
CG	Conjugate Gradients
CIG	Common Image Gather
CMP	Common MidPoint gather
CSG	Common Shot Gather
DMO	Dip MoveOut
DSR	Double Square Root

FFD Fourier Finite Difference

FFT Fast Fourier Transform

IFFT Inverse Fast Fourier Transform

IRLS Iterative Reweighted Least-Squares

NMO Normal MoveOut

PCG Preconditioned Conjugate Gradients

PLSM Preconditioned Least-Squares Migration

PSPI Phase Shift Plus Interpolation

RLSM Regularized Least-Squares Migration

RTT Radial Trace Transform

SLSM Sparse Least-Squares Migration

SSR Single Square Root

Chapter 1

Introduction

Seismic data contain information about traveltimes and waveform amplitudes. The former is tightly related to geological structures, and the latter is controlled by elastic rock properties. In structural exploration, geophysicists mainly focus on the structural imaging problem to find favorable structures. However these structures do not always contain hydrocarbons. More information should be used to decrease the risk of oil/gas exploration. For this reason, efforts have been made to use seismic amplitudes to infer rock properties and/or hydrocarbon indicators. For P-wave (compressional wave) exploration, the angle-dependent reflectivity can be expressed explicitly in terms of three rock parameters, compressional wave velocity perturbation, shear wave velocity perturbation and density perturbation (Aki and Richards, 1980). With redundant angle-dependent amplitude arising from multi-channel seismic data, we can use Zoeppritz equations (1980) to invert the aforementioned media parameters. Since the existence of oil and gas could significantly change the media parameters and consequently introduce anomalies on the angle-dependent amplitudes, such anomalies have become an important tool as direct hydrocarbon indicators. The success of these techniques rely on high-quality angle-dependent amplitudes, efficient extraction of media properties from amplitude information and good attribute-analysis strategies.

Considering a simple flat interface, an incidence P wave (compressional wave) on one side will result in both S wave and P wave responses on the other side in the form of reflected energy and transmitted energy (see Figure 1.1). The reflectivity of the $\grave{P}\acute{P}$ wave (the reflectivity in respect to downgoing P wave and upgoing P wave) can be approxi-

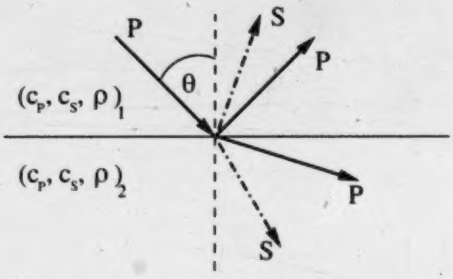


Figure 1.1: Incident P wave and the reflected and transmitted waves at an interface between two media. P: compressional wave. S: shear wave. θ: the incident angle. c_P: P wave velocity. c_S: shear wave velocity. ρ: density. The subscript 1 and 2 denote the upper medium and the lower medium, respectively.

mated by the following formula (Aki and Richards, 1980)

$$R_{\grave{P}\acute{P}}(\theta) \approx \frac{1}{2}(1 - 4\frac{V_s^2}{V_p^2}\sin^2\theta)\frac{\Delta\rho}{\rho} + \frac{\sec^2\theta}{2}\frac{\Delta V_p}{V_p} - \frac{4V_s^2}{V_p^2}\sin^2\theta\frac{\Delta V_s}{V_s} \tag{1.1}$$

where

$$\begin{aligned} \Delta V_p &= (V_{p2} - V_{p1}), \\ V_p &= (V_{p1} + V_{p2})/2, \\ \Delta V_s &= (V_{s2} - V_{s1}), \\ V_s &= (V_{s1} + V_{s2})/2, \\ \Delta\rho &= (\rho_2 - \rho_1), \\ \rho &= (\rho_2 + \rho_1)/2, \end{aligned}$$

where the subscript 1 denotes the medium above the interface, and subscript 2 denotes the medium bellow the interface. The task of AVA analysis by $\grave{P}\acute{P}$ wave is to invert the rock properties: density perturbation, P wave velocity perturbation and S wave velocity perturbation, given multiple measurements of angle-dependent reflectivity.

This thesis focuses on providing accurate angle-dependent amplitude ('true amplitude') using inverse theory and wave-equation based operators. The concept 'true amplitude' means all recovered amplitudes are in proportion to the true reflection response of the earth model.

Conventional AVA (Angle Versus Amplitude) analysis uses data that are preprocessed by Normal Moveout (NMO), an operator correcting far offset traveltimes to match the near offset traveltimes. This operator assumes flat layers. In NMO processing, seismic data are sorted in CMP (Common Midpoint) gathers, and each CMP gather is processed independently. The goal of NMO correction is to align seismic events to their normal incident (zero-offset) traveltimes, which makes it easier to pick the amplitude along the offset direction. After extracting the amplitude of different offsets, we can convert AVO (Amplitude Variation with Offset) to AVA (Amplitude Variation with Angle) by ray-tracing methods. Figure 1.2 displays the basic NMO procedure that produces an AVO panel with only a flat event.

The shortcoming of NMO method is obvious. First, the assumption of flat layers is violated in most cases, which means that the traveltime correction and ray-tracing estimation of incident angles is inaccurate. For this reason, NMO can not correctly locate dipping events and complex structures. Second, the application of this method is hampered by NMO stretching introduced in the wavelet and tuning effects[1] caused by thin layers.

For dipping events, the accuracy of travel time correction can be increased by Dip Moveout (DMO) correction. Hale (1984) developed a Fourier DMO algorithm to preserve the amplitude of events with conflicting dips. The method uses constant-offset gathers and transports energy across midpoints. Therefore it can be regarded as a partial migration. More accurate DMO methods with geometry calibrations were proposed by Notfors et al. (1987), Liner, C. L. (1990) and Black et al. (1993). Zhou et al. (1996) presented an efficient algorithm by implementing Black's DMO algorithm in the log-stretch $F-K^2$ domain. DMO theory is based on a simple media, one upper layer and one subspace, which is improper since usually the earth model has more than two layers. Therefore the DMO processed result can also be unsatisfactory. On the other hand, migration/inversion can better address the problem by solving the wave-equation in various ways.

The migration of P wave seismic data is to solve zero time wavefield $(p(x, y, z, t = 0))$

---

[1] Overlapping of neighboring events

[2] Frequency-wavenumber domain

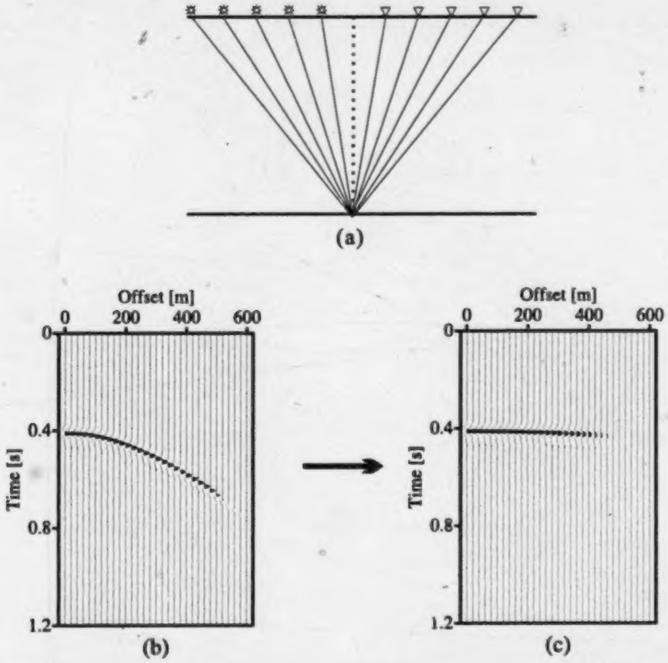


Figure 1.2: A simple example of NMO correction. (a) The geometry for the data acquisition. The data were produced by many experiments by burying sources (symbolized in stars) to the left of the midpoint at different offsets. The receivers (symbolized in triangles) were buried at the other side. The lines connecting sources and receivers are the ray paths of the seismic experiments. (b) The relevant Common Midpoint (CMP) gather. (c) The CMP gather after NMO correction.

from the following acoustic approximate wave equation:

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}+(\frac{\partial^2 p}{\partial x^2}+\frac{\partial^2 p}{\partial y^2}+\frac{\partial^2 p}{\partial z^2})=0, \tag{1.2}$$

where p denotes 3-D seismic response ($p = p(x, y, z, t)$), c denotes 3-D velocity model ($c = c(x, y, z)$), given $p(x, y, z = 0, t)$, i.e., migration reconstructs the initial wavefield in 3-D volume from the observed time-lapse seismic response at the surface.

There are two kinds of methods to solve the above problem. One is based on ray and scattering theory, called Kirchhoff migration. This kind of methods use high-frequency approximation and solve amplitude and phase separately (Bleistein et al., 2000). The amplitude can be improved by an asymptotic inversion (Bleistein and Gray, 2002) or the least-squares method (Nemeth et al., 1999; Xu et al., 1998). The accuracy of Kirchhoff style methods is limited by the high-frequency approximation and the effects of multi arrivals. The other kind of methods are called wave-equation migration that directly solves the wave equation with numerical implementation. Among them are finite difference migration (Claerbout, 1985; Mufti et al., 1996), Fourier finite difference migration (Ristow and Rühl, 1994), Fourier migration (Stolt, 1978) and phase shift migration (Gazdag, 1978; Gazdag, 1984). Generally, wave-equation migration is more accurate than Kirchhoff migration since no high-frequency approximation is involved, and multi arrivals are automatically considered.

No matter what method is used, we can assume the earth as a linear system. This is justified by removing multiples from the original seismic data. De-multiple is an important preprocessing before inverting the seismic data for structural and petrophysical features. However the linearization of the system may be understood in different ways as of the involved operators. In Kirchhoff methods, as explained in Chapter 2, linearization means truncating higher orders of the expansions series of the traveltimes and amplitudes. In wave equation based methods, the coupling between upgoing waves and downgoing waves is ignored, which allows us to solve the upgoing waves in a linear way. In this thesis, I am interested in implementing wave equation based methods for structure and amplitude studies.

Wave-equation migration can be applied with different strategies considering the acquisition geometry. Shot-profile migration is more suitable for wide-azimuth land data, and shot-geophone (DSR) migration is good for marine data (Biondi, 2003b). Theoretically, it has been proved that these two methods will produce equivalent results (Berkhout, 1982; Wapenaar et al., 1987; Biondi, 2003b). However we should take special care of the aperture for shot-profile migration (Huang et al., 2003). Therefore, in this thesis I use DSR operators for migration/inversion.

In wave-equation migration, there are basically two steps toward AVA imaging. First we can use any method to extrapolate the wavefield to different depths. Therefore the data are kept in a prestack data volume (Mosher et al., 1996). The second step is to apply the imaging condition to produce local angle/ray-parameter image gathers. For shot-profile migration, de Bruin, et al.(1990) use local slant stack to convert offset domain data to ray-parameter domain image. Similarly, this technique is used for shot-geophone (DSR) migration (Prucha et al., 1999). The ray-parameter domain image gathers can be converted to angle-domain gather by a simple formulation (Prucha et al., 1999; Fomel, 2004). However the formula is dependent on structural dip and interval velocity, which can be often inconvenient.

One alternative way of producing angle-domain image gathers is to apply a transform in the offset image domain. Sava and Fomel (2003) proposed a method directly converting image gathers from offset domain to angle domain. The method is independent of structural dip and interval velocity. However it is sensitive to migration velocity mismatch. In my least-squares migration implementation, I use Prucha's method since it is less sensitive to velocity errors.

The described two steps of wave-equation migration are the conventional routine of imaging processing for AVA analysis. It is better than NMO and DMO at handling complex geological structure. However, the seismic data deserve further application. In the viewpoint of inversion, the conventional migration is called the adjoint of the modelling procedure, which only provides an approximate solution. One drawback of this method is that the produced result does not fit the data. Although the solution can provide a good framework for structural imaging, it does not produce good amplitudes. An economical remedy is to apply some corrections (the inverse of the imaging Jacobian) to the wavefield before the imaging condition is applied. However the application of this approach is hampered by the data quality and the complexity of the earth model. First, the data can be irregularly sampled, which brings aliasing in the solution. Second, when the velocity variation is too large, the assumption of the imaging Jacobian theory is violated. On the other hand, iterative inversion methods can efficiently address these problems.

Posing migration as an inverse problem and solving it by iterative methods have three major advantages. First by inverting a linear system, the seismic observations are properly honored. The solution is optimized to be accurate both kinematically and dynamically. Second, efficient iterative algorithm like Conjugate Gradients (CG) (Hestenes and Steifel, 1952) can be used, which avoids the need of assembling large matrices for the inversion. Third, *a priori* information of the data and the model can be utilized to regularize the inversion. For example, smooth regularization has been used to develop robust least-squares migration (Ehinger and Lailly, 1991; Kuehl and Sacchi, 1999; Prucha and Biondi, 2002; Kuehl and Sacchi, 2003). The geometry of data acquisition can be simulated by a sampling matrix (Duquet et al., 2000; Kuehl and Sacchi, 1999; Prucha and Biondi, 2002; Kuehl and Sacchi, 2003). Inverting this matrix helps to alleviate the artifacts due to data acquisition footprints.

This thesis is organized as follows. In Chapter 2, I review two styles of seismic modelling and the adjoint operators. The Kirchhoff method is briefly discussed and used to exemplify the concept of migration and inversion. On the other hand, the wave equation based operators are given in full detail. The derivation of wave-equation AVP modelling/migration starts from a review of two-way and one-way representation of the acoustic wave equation. Then the one-way wave-equation formula is implemented into shot-profile modelling/ migration. The shot-profile and DSR continuation methods are unified by analyzing the dispersion relationship of the two methods. The equivalence of these two methods are confirmed by a toy example of AVP imaging problem. The formulas of common-azimuth modelling/ migration are included, which are the fundamental operators of this thesis to address 3-D common-azimuth data. In Chapter 3, I generalize the theory of least-squares migration with various tastes of regularization. The problems are solved by the CG algorithm in an iterative way. To test the inversion method, I present two simple examples. One is the inversion of a smooth model from a noisy dataset, which is used to examine the robustness of smooth regularization. The other example is a multi-channel deconvolution problem to test the idea of combining smooth regularization and sparse regularization. The former regularization is forced in the ray parameter direction, and the latter is applied in the depth direction. A routine of Iterative Reweighted Least-Squares (IRLS) (Scales and Smith, 1994)is described in detail to solve the sparsely-regularized problem. The choice of hyperparameters is examined by a test of sparse deconvolution. In Chapter 4 and Chapter 5, some synthetic and field data are used to test the theory of migration and regularized inversion. Finally, major contributions of this thesis are provided in the discussions and conclusions. Difficulties and future directions are also discussed.

Chapter 2

Seismic modelling and migration/the adjoint

2.1 Introduction

Seismic modelling is a forward problem. In the geophysical community it is used to simulate seismic survey from an earth model, which can be expressed in terms of rock properties (velocities, densities, etc) or reflectivities. Oppositely, the migration/inversion is an inverse problem. Given observations at the surface of the earth, we try to reconstruct the earth model. Although modelling and migration are different, they are connected by the theory of wave propagation. For simplicity the wavefield can be represented by an acoustic wave equation. Both modelling and migration involve solving the wave equation but with different conditions. The modelling problem is an initial condition problem that predicts the seismic response given properties of the media and the source. The migration problem is a boundary condition problem in the sense that we know the temporal seismic response at the surface. The product of migration is an image or image gathers representing a scaled reflectivity model of the subsurface. This model can be afterwards used to invert for rock properties and/or to estimate hydrocarbon indicators. Conventional migration is usually not the real inverse operator of the forward problem but the so-called adjoint operator. If we use the model produced by the adjoint operator for modelling, the output data can not fit the observed data. For this reason, I would rather regard inversion as the procedure to find a solution honoring the data. Broader views of migration/inversion are discussed by some authors (Weglein and Stolt, 1999;

Bleistein et al., 2000).

In chapter 3, I will show that modelling and migration/the adjoint are essential operators for linear and non-linear inversion. Before moving onto the application of inverse theory, it is necessary to clarify these operators. The choice of modelling and the adjoint is fundamental to the accuracy of inversion. Therefore, in this chapter I first explain the concept of seismic modelling and migration by the Kirchhoff approach, which is most frequently used in the industry. I discuss the advantage and disadvantage of this method, which motivates the use of wave equation based techniques. A comparison of migration and inversion is previewed by an experiment with 1-D seismic data. Then I will provide a detailed study of the wave-equation approach, including the basic theory of two-way/one-way wave equation modelling/migration, shot profile vs. DSR operator, common-azimuth approximation, angle-domain imaging condition and velocity correction techniques to handle complex earth models with strong lateral velocity variations.

2.2 Kirchhoff modelling and migration/the adjoint

2.2.1 WKBJ approximation

Kirchhoff method is based on ray theory, which solves the wave equation by decomposing it into two equations, eikonal equation and transport equation, for travel time and amplitude respectively. Here the words 'solves wave equation' mean that we can analytically express the wavefield in terms of some kernels of travel time and amplitude. Or we can say that the solution given by ray theory represents the wavefield in a more explicit way than the original wave equation. The representation allows us to simulate seismic response by some transformation of the earth model that we are interested in. With further approximation (Kirchhoff approximation), we can derive a formula that links seismic response to a reflectivity model. In the Kirchhoff theory, we assume that the wavefield can be represented in the following form (WKBJ approximation):

$$u(\mathbf{x},\omega) \sim \omega^{\beta} e^{i\omega\tau(\mathbf{x})} \sum_{j=0}^{\infty} \frac{A_j(\mathbf{x})}{(i\omega)^j}, \tag{2.1}$$

where $\mathbf{x}$ denotes the 3-D location, ω is the temporal frequency, A_j is the j th order amplitude and β is an integer number. In the temporal frequency and space domain, the wave equation can be expressed as follow:

$$Lu = \left[\nabla^2 + \frac{\omega^2}{c^2(\mathbf{x})}\right]u(\mathbf{x},\omega) = 0. \tag{2.2}$$

Substituting a trial solution in the WKBJ form in equation 2.2 leads to (Bleistein et al., 2000):

$$Lu = \omega^\beta e^{i\omega\tau}\sum_{j=0}^{\infty}\frac{1}{(i\omega)^j}\left[\omega^2\left\{\frac{1}{c^2} - (\nabla\tau)^2\right\}A_j + i\omega\{2\nabla\tau\cdot\nabla A_j + A_j\nabla^2\tau\} + \nabla^2 A_j\right] = 0. \tag{2.3}$$

Usually different powers of the temporal frequency ω can not cancel each other. So by setting the first two terms to zeros, we have the following two equations:

$$(\nabla\tau(\mathbf{x}))^2 - \frac{1}{c^2(\mathbf{x})} = 0, \tag{2.4}$$

$$2\nabla\tau(\mathbf{x})\cdot\nabla A_0(\mathbf{x}) + A_0\nabla^2\tau(\mathbf{x}) = 0, \tag{2.5}$$

where τ is the travel time, and A_0 is the leading order amplitude. Equation 2.4 is called the eikonal equation, and equation 2.5 is called the transport equation. Higher order amplitude terms can be calculated by recursively solving

$$2\nabla\tau\cdot\nabla A_j + A_j\nabla^2\tau = -\nabla^2 A_{j-1}, \tag{2.6}$$

where j and $j-1$ are order numbers of the amplitude term ($j > 0$). In imaging applications only the leading order transport equation is required to represent the amplitude.

The eikonal equation 2.4 can be solved by the method of characteristics (Bleistein et al., 2000). This method converts the partial derivative equation to a sets of ordinary equations. The ordinary equations may have various forms depending parameterization. One often used pair of equations are listed as follows

$$\frac{d\mathbf{x}}{d\tau} = c^2(\mathbf{x})\mathbf{p}, \quad \frac{d\mathbf{p}}{d\tau} = -\frac{\nabla c(\mathbf{x})}{c(\mathbf{x})}, \tag{2.7}$$

where $\mathbf{p}$ is the vector of wavefield gradient ($\mathbf{p}= (u_{x1}, u_{x2}, u_{x3})$), τ is the travel time at location $\mathbf{x}$, and $c(\mathbf{x})$ is the wave speed. These equations can be solved by raytracing methods.

The transport equation can be solved by dynamic raytracing (Cerveny, 2000). Alternatively, an economical way of calculating amplitude is to use a geometrical spreading factor by assuming constant velocity (Aki and Richards, 1980):

$$A \propto \frac{1}{|\mathbf{x} - \mathbf{x}_0|} \tag{2.8}$$

for 3-D case and

$$A \propto \sqrt{\frac{1}{|\mathbf{x} - \mathbf{x}_0|}} \tag{2.9}$$

for 2-D case, where $\mathbf{x}$ symbolize the location where we measure the amplitude and $\mathbf{x}_0$ is the source position.

The above geometrical spreading formulas can be modified by changing the distance to the accumulated length of the ray path, which is available from the travel time raytracing routine. In general, we assume all amplitude have a same scale with respect to the geometrical spreading factor, therefore for simplicity we can replace "$\propto$" with "=" in the above relationship.

2.2.2 Green's function for linear partial differential operators and derivation of Kirchhoff modelling

The Green function is a integral kernel that can be used to solve differential equations. It can be defined in a general way as follows (Arfken, 1985). Given a linear partial differential operator L, if we can find a function G such that

$$LG(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \tag{2.10}$$

then the solution to $Lu = f$ is

$$u(\mathbf{r}) = \int G(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') d^3\mathbf{r}', \tag{2.11}$$

and G is called the Green's function of the linear operator L.

For the acoustic wave equation, the operator L is replaced with the Laplacian operator ∇. Given a macro velocity model, we can calculate the Green's function by solving for amplitude and travel time terms separately as previously described. The wave equation can be formulated in the temporal frequency and space domain as follows:

$$Lu(\mathbf{x}, \mathbf{x}_s, \omega) = -F(\omega)\delta(\mathbf{x} - \mathbf{x}_s) \tag{2.12}$$

The total wave field u can be decomposed into two parts, the unperturbed incident wave u_i and the perturbed wavefield or the reflection u_s. The incident wave is closely related to the source, therefore we decompose equation 2.12 into two equations:

$$Lu_I = -F(\omega)\delta(\mathbf{x} - \mathbf{x}_s) \tag{2.13}$$

$$Lu_s(\mathbf{x}, \mathbf{x}_s, \omega) = 0 \tag{2.14}$$

A Green function is defined for the receiver $\mathbf{x}_g$ by

$$LG(\mathbf{x}, \mathbf{x}_g, \omega) = -\delta(\mathbf{x} - \mathbf{x}_g) \tag{2.15}$$

Now we apply Green's theorem to G and u.

$$\int_{D_a} (GLu - uLG)dv = \int_{S_a+S_R} (G\frac{\partial u}{\partial n} - u\frac{\partial G}{\partial n})ds, \tag{2.16}$$

where D_a is the upper hemisphere volume shown in Figure 2.1, S_a is the upper surface of the sphere, and S_R is the enclosed part of S by the surface of the sphere.

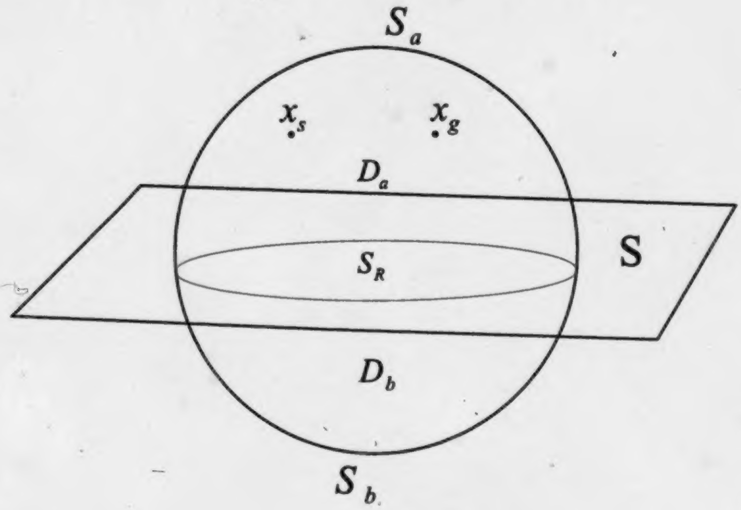


Figure 2.1: Diagram for deriving the Kirchhoff integral operator (Bleistein et al., 2000).

To continue the derivation, we can calculate the left side of equation 2.16 as the following:

$$\int\limits_{D_a} GLu = \int\limits_{D_a} G(\mathbf{x},\mathbf{x}_g,\omega)(-F(\omega)\cdot\delta(\mathbf{x}-\mathbf{x}_s))dV = -G(\mathbf{x}_s,\mathbf{x}_g,\omega)\cdot F(\omega), \tag{2.17}$$

and

$$\int\limits_{D_a} -uLGdv = \int\limits_{D_a} -u(\mathbf{x},\mathbf{x}_s,\omega)(-delta(\mathbf{x}-\mathbf{x}_g))dV = u(\mathbf{x}_g,\mathbf{x}_s,\omega) \tag{2.18}$$

Using 2.17 and 2.18 in 2.16, we have

$$\begin{aligned} u(\mathbf{x}_g,\mathbf{x}_s,\omega) - F(\omega)\cdot G(\mathbf{x}_s,\mathbf{x}_g,\omega) &= \int\limits_{S_a+S_R} (G\frac{\partial u}{\partial n} - u\frac{\partial G}{\partial n})dS \\ u_I + u_s - F(\omega)\cdot G(\mathbf{x}_s,\mathbf{x}_g,\omega) &= \int\limits_{S_a+S_R} (G\frac{\partial u}{\partial n} - u\frac{\partial G}{\partial n})dS \end{aligned} \tag{2.19}$$

Similar equation can be derived by applying Green's theorem to u_I and G:

$$u_I(\mathbf{x}_g,\mathbf{x}_s,\omega) - F(\omega)G(\mathbf{x}_s,\mathbf{x}_g,\omega) = \int\limits_{S_R+S_a} (G\frac{\partial u_I}{\partial n} - u_I\frac{\partial G}{\partial n})dS \tag{2.20}$$

Combining 2.19 and 2.20, we find

$$u_s(\mathbf{x}_g,\mathbf{x}_s,\omega) = \int\limits_{S_R+S_a} (G\frac{\partial u_s}{\partial n} - u_s\frac{\partial G}{\partial n})dS \tag{2.21}$$

According to Somerfield condition, the integral over the surface S_a should be zero when the sphere is infinite in size ($R \rightarrow \infty$). Therefore, 2.21 becomes

$$u_s(\mathbf{x}_g,\mathbf{x}_s,\omega) = \int\limits_{S} (G\frac{\partial u_s}{\partial n} - u_s\frac{\partial G}{\partial n})dS \tag{2.22}$$

By adopting the Kirchhoff approximation $u_S = Ru_I$, the above equation can be changed to

$$u_s(\mathbf{x}_g,\mathbf{x}_s,\omega) = \int\limits_{S} R(u_I\frac{\partial G}{\partial n} + G\frac{\partial u_I}{\partial n})dS\ , \tag{2.23}$$

where R is the angle-dependent reflectivity.

From 2.13 we know the incident wave can expressed in the following form:

$$u_I = F(\omega)A(\mathbf{x}, \mathbf{x}_s, \omega) \cdot e^{i\omega\tau(\mathbf{x},\mathbf{x}_s)} \tag{2.24}$$

The green's function of the receiver, $G(\mathbf{x}, \mathbf{x}_g, \omega)$, is in the following form:

$$G(\mathbf{x}, \mathbf{x}_g, \omega) = A(\mathbf{x}_g, \mathbf{x})e^{i\omega\tau(\mathbf{x},\mathbf{x}_g)} . \tag{2.25}$$

Substituting 2.24 and 2.25 in 2.23 leads to

$$\begin{aligned} u_s(\mathbf{x}_g, \mathbf{x}_s, \omega) \quad &= i\omega F(\omega) \int_S RA(\mathbf{x}_g, \mathbf{x})A(\mathbf{x}, \mathbf{x}_s) \cdot (\mathbf{n} \cdot \nabla[\tau(\mathbf{x}_g, \mathbf{x}) + \tau(\mathbf{x}, \mathbf{x}_s)]) \cdot \\ &e^{i\omega(\tau(\mathbf{x}_g,\mathbf{x})+\tau(\mathbf{x},\mathbf{x}_s))} \cdot dS. \end{aligned} \tag{2.26}$$

Equation 2.26 can be further simplified into the following general form (Bleistein and Gray, 2002):

$$u_s(\mathbf{x}_g, \mathbf{x}_s, \omega) = S(\omega) \int R(\mathbf{x}, \theta)A(\mathbf{x}_g, \mathbf{x}, \mathbf{x}_s)|\frac{2\cos\theta}{c(\mathbf{x})}|e^{i\omega\tau(\mathbf{x}_s,\mathbf{x},\mathbf{x}_g)}dS, \tag{2.27}$$

where

$$S(\omega) = \begin{cases} -i\omega F(\omega) & \text{for 3D} \\ |\omega|F(\omega) & \text{for 2D} \\ \sqrt{|\omega|}\sqrt{\frac{\sigma_s\sigma_r}{\sigma_s+\sigma_r}}e^{3i\pi sgn(\omega)/4}F(\omega) & \text{for 2.5D,} \end{cases}$$

$$A(\mathbf{x}_g, \mathbf{x}, \mathbf{x}_s) = A(\mathbf{x}, \mathbf{x}_s)A(\mathbf{x}_g, \mathbf{x})$$

$$\tau(\mathbf{x}_s, \mathbf{x}, \mathbf{x}_g) = \tau(\mathbf{x}_s, \mathbf{x}) + \tau(\mathbf{x}, \mathbf{x}_g)$$

The symbol θ is the reflection angle. By far, we have derived a forward modelling scheme that links the model, angle-dependent reflectivity, with the seismic data. The Kirchhoff modelling involves linear summation of wave perturbation throughout an earth subvolume. For a source and receiver pair, we need to calculate the Green functions of all combinations of source ($\mathbf{x}_s$), imaging point ($\mathbf{x}$) and the receiver ($\mathbf{x}_g$). For each combination, we can calculate the contribution from this ray path using the reflectivity model and the Green function. Then the contributions of all rays are summed up to simulate the seismic response. Oppositely, inversion reconstructs the reflectivity model from the seismic data, given a background velocity model. Similar to modelling, inversion also needs Green functions.

There are two kinds of methods to solve the inverse problem: direct and indirect methods. The direct method attempts to transform data to model by using a formula. Usually some approximation theory is involved in such manipulation (Bleistein and Cohen, 1979; Cohen and Bleistein, 1979; Clayton and Stolt, 1981; Beylkin, 1985; Miller et al., 1987; Beylkin and Burridge, 1990). These methods are attractive since they are more economical than the second way of inversion: iterative optimization methods. However, direct methods requires high-quality data (Bleistein et al., 2000) and may lead to instabilities (Jin et al., 1992). Especially when data are incomplete, the solution can be unsatisfactory. The second type of methods, iterative methods, involve fitting the seismic data by recursively switching between data space and model space to seek a solution that can fit the data. The advantage of iterative methods are obvious: the solution is constructed to honor the data. Furthermore, *a* priori information can be used in the inversion. For example smoothness across the midpoints and the quality of the acquired data can be incorporated as model and data space prior information, respectively (Duquet et al., 2000). The shortcoming of iterative methods is that they can be quite demanding from the computational point of view. However, this will not be an an issue with large high-performance computing facilities being more and more accessible. Moreover, as shown in this thesis, it is possible to acquire promising results in only a few of iterations by proper preconditioning of the problem.

2.2.3 Migration/the adjoint and a preview of inversion

A stable and often practice of Kirchhoff imaging is to apply the adjoint of the modelling operator to seismic data. It can be formulated as follows:

$$R(\mathbf{x},\theta) = \int S^*(\omega)A(\mathbf{x}_g,\mathbf{x},\mathbf{x}_s)|\frac{2\cos\theta}{c(\mathbf{x})}|e^{-i\omega\tau(\mathbf{x}_g,\mathbf{x},\mathbf{x}_s)}d\omega, \qquad (2.28)$$

where $S^*(\omega)$ is the adjoint of $S(\omega)$ in 2.27. This method is cheaper than the aforementioned direct inversion, but the image is often smeared and the amplitude information is distorted. The output is usually only suitable for structural studies. However by modelling and the adjoint of modelling, we can develop an inversion scheme to fit the seismic data and provide more accurate solution, both kinematically (for structures) and dynamically (for amplitudes).

To clarify the above assertion, I present an example of 1-D Kirchhoff migration/ inversion. The modelling operator is a simplified version of equation 2.27 by setting θ to zero. I prepare a model with constant velocity $c = 2000m/s$. The forward modelling operator is

$$u_s(\mathbf{x}_0, \omega) = \omega F(\omega) \int R(\mathbf{x}) \cdot A \cdot \frac{2}{c} \cdot e^{i\omega\tau} d\mathbf{x}, \tag{2.29}$$

where $\mathbf{x}_0 = \mathbf{x}_g = \mathbf{x}_s$(zero-offset geometry), $A = \sqrt{(1/|\mathbf{x} - \mathbf{x}_0|)}$ and $\tau = 2 * |\mathbf{x} - \mathbf{x}_0|/c$. The adjoint operator is

$$R(\mathbf{x}) = \int \omega F^*(\omega) \cdot u_s(\mathbf{x}_0, \omega) \cdot \frac{2}{c} \cdot e^{-i\omega\tau} d\omega. \tag{2.30}$$

Please note that in equation 2.30 the variable $\mathbf{x}$ is hidden in $\tau(\mathbf{x})$. Five reflectors are placed at z = 200 m, 350 m, 415 m, 550 m and 700 m respectively. The synthetic seismic trace is shown in Figure 2.2b.

Figure 2.2c shows that the adjoint operator provides a smeared version of the model, and the effects of the wavelet is obvious. The reflector position is correctly placed at the peak of the waveform, but the amplitude is far from that of the true model. On the other hand, the wavelet is suppressed in the iteratively inverted model (see Figure 2.2d), and the reflectivity is in proportion to the true value. Some spurious sidelobes are still present around the spikes. Such artifacts can be further suppressed by sparse regularization of the model, which will be discussed in future chapters.

2.3 Some comments on the Kirchhoff operator

Kirchhoff method will still be used for a while due to two practical reasons. First, the method can be implemented for economical inversion. The most difficult part is the cal-

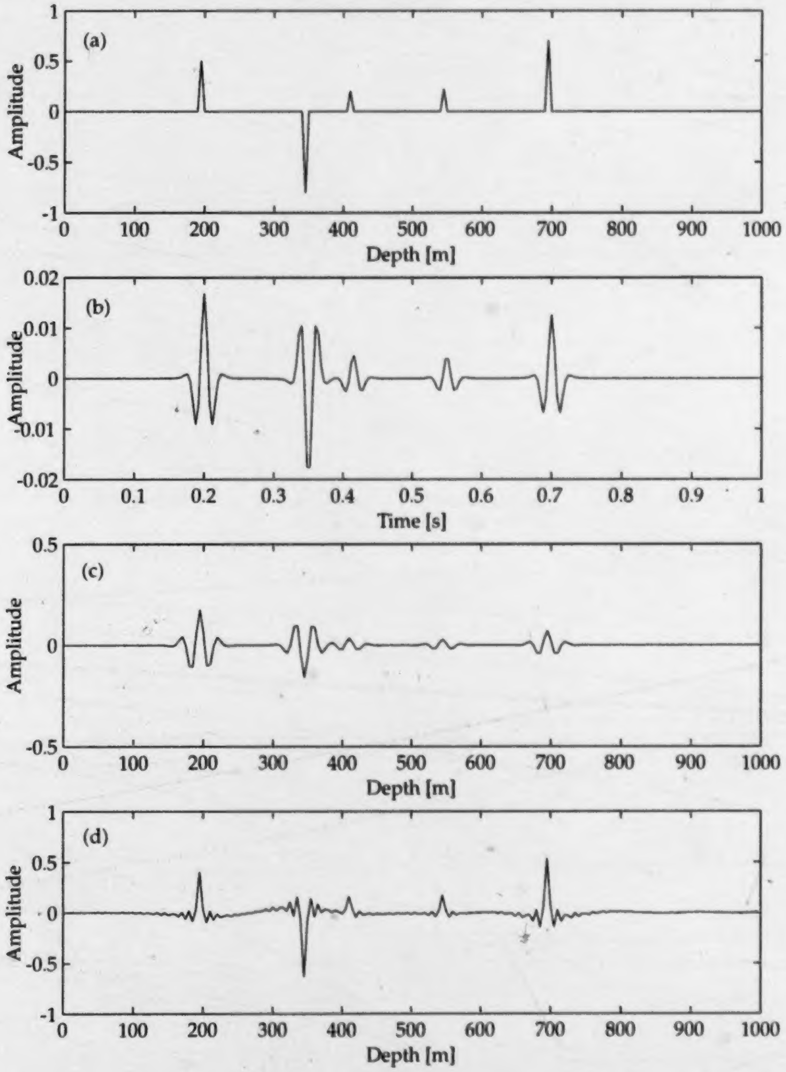
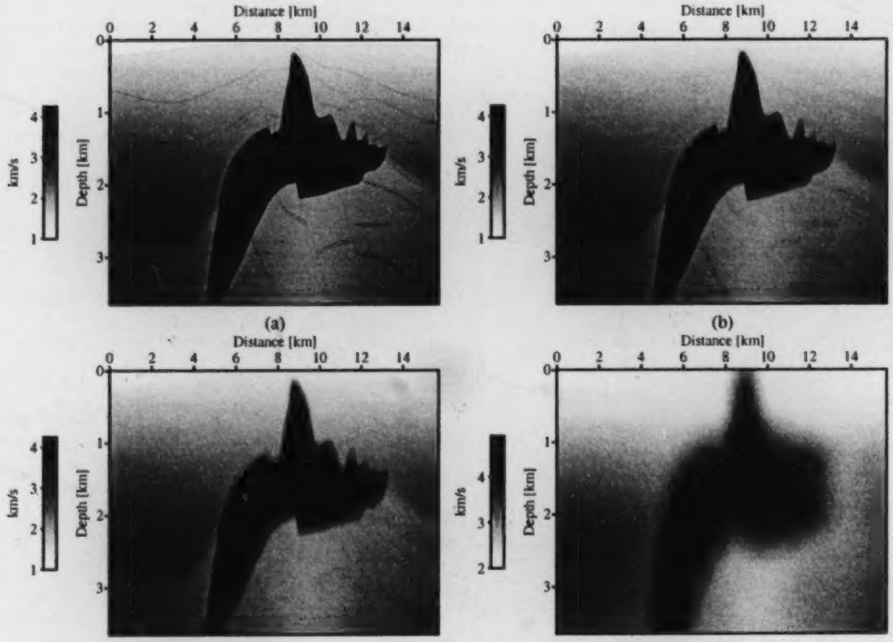


Figure 2.2: An example of 1-D Kirchhoff inversion. (a) The true reflectivity model. (b) The synthetic seismic trace with a sampling rate of 0.004s. (c) The calculated reflectivity model by the adjoint operator. (d) The reconstructed reflectivity model by a linear iterative inversion.

culation of Green's functions, which involves raytracing for travel times and amplitudes. Once these functions have been calculated, we can use them repeatedly. Kirchhoff modelling and migration are economical since only linear summations are needed. The cost can be further decreased by controlling the summation aperture (Biondi, 2003a). Second, Kirchhoff method can be operated in a target-oriented mode, and it is flexible in handling irregular data acquisition.



(a) (b) (c) (d)

Figure 2.3: Velocity profile A-A' from the SEG/EAGE data. (a) Original model. (b) Smoothed model (square width=5). (c) Smoothed model (square width=10) . (d) Smoothed model (square width=50).

The shortcomings of Kirchhoff method are also obvious. First, the method is based on high-frequency approximation, which requires large distance between diffractors and sources or receivers. This makes near-surface image inaccurate (Gray et al., 2001). Second, more than one raypath may exist between the source and the receiver, which causes

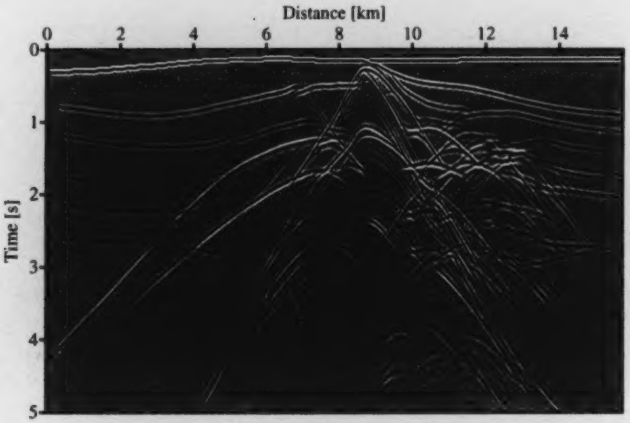
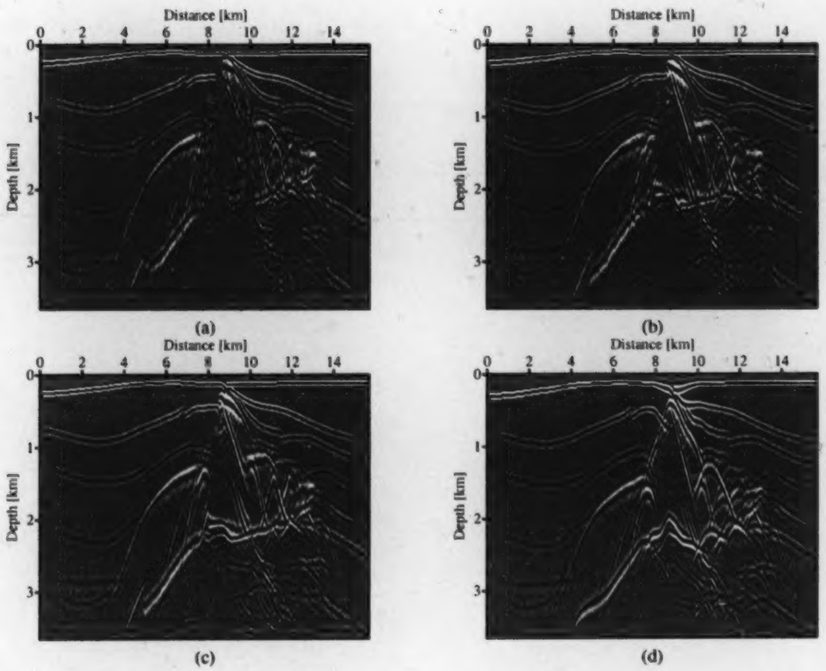


Figure 2.4: Zero-offset data for the profile A-A′ from SEG/EAGE model.

confusion during data summation. Especially, when the structure is complicated, the method will fail to accurately show structural details. The multipathing problem can be alleviated by some local Kirchhoff methods (Hill, 1990; Hill, 2001; Bevc, 1997), but the image quality is generally not so good as that obtained via wave-equation methods like split-step (Popovici, 1996) or PSPI (Gazdag, 1984).

To end this section, I evaluate the performance of Kirchhoff method by a benchmark data set: poststack 2-D SEG/EAGE data set. The model contains some faults and a big salt body. It is a challenge to illuminate the subsalt reflectors and accurately resolve the salt body. The zero-offset data were produced by an exploding reflector modelling method using the velocity model shown in Figure 2.3a. The density is set to a constant. The synthetic seismic section (Figure 2.4) includes the signature of all reflectors, but direct interpretation of the data is difficult without migration.

Before applying Kirchhoff migration, usually the velocity model is smoothed to estimate the macro velocity field. No one can tell that what kind of smoothing method is perfect, and it is interesting that even when the macro velocity is far from the real model, the Kirchhoff method can still produce an approximate solution. In this test, I use the average velocity of points within a square to represent the macro velocity at the center of that square. The larger is the square, the smoother is the calculated macro velocity model. Figure 2.3 b, c, and d show the smoothed velocity models by using squares with widths of 5, 10, 50, respectively. The migrated images 2.5 show that both under smoothing and over smoothing of the velocity field give inferior results. In all cases, the subsalt reflections cannot be accurately mapped in the image. The edge of the salt body is not satisfactory either. This is often a problem when we use Kirchhoff methods to image complex media. Please note that a better result of this data set has been shown by O'Brien and Gray (1996). However, as that result shows, complex structure imposes a challenge to Kirchhoff imaging methods. The point here is to show the difficulty of preparing velocity model for ray tracing that is required by the algorithm. For this reason, I prefer wave equation operators, which are discussed in the next few sections. Wave-equation migration methods usually need more accurate velocity information, and they are sensitive to velocity mismatch, which makes these methods suitable for velocity analysis.



(a) (b) (c) (d)

Figure 2.5: Migrated images by Kirchhoff method using different macro velocity models. (a) Result using the true velocity model. (b) Result using velocity model shown in Figure 2.3b. (c) Result using velocity model shown in Figure 2.3c. (d) Result using velocity model shown in Figure 2.3d.

2.4 Wave-equation modelling/migration

Wave equation based methods are derived in a more formal way than Kirchhoff methods. In this kind of methods, amplitude and travel time is solved at the same time in contrast to the Kirchhoff style approach of solving the decomposed eikonal equation and transport equation. No high-frequency approximation is involved. Furthermore, multi-pathing problem is implicitly solved. For these reasons, wave-equation methods usually provide better results than ray-based methods especially for complex media.

2.4.1 Two-way versus one-way wave-equation representation

In a seismic experiment, an impulse is triggered at the source and the wave propagates down to the sub-volume. When the wavefront hits a reflector (an interface between two contacting media), some part of the energy transmits through the interface and continues its journey to the deeper area. Some part is reflected back toward the surface. Therefore, we can think the total wavefield as a combination of downgoing wave and upgoing wave. These two waves may couple each other at high angles in inhomogeneous media (Wapenaar and Berkhout, 1985). Depending on the way of treating these waves, two-way and one-way wave-equation methods were proposed (Wapenaar and Berkhout, 1985; Wapenaar and Berkhout, 1986a; Wapenaar and Berkhout, 1986b).

By ignoring the source, the equation of motion can be formulated in the temporal frequency domain as follows:

$$\nabla P = -i\omega\rho\mathbf{V}, \tag{2.31}$$

where P is the pressure, ω is the temporal frequency and $\mathbf{V}$ is the velocity vector. The stress-displacement equation reads

$$K\nabla \cdot \mathbf{V} = -i\omega P, \tag{2.32}$$

where K is the bulk compression modulus. For acoustic media, $K = c^2 \cdot \rho$. From equation

2.31 we have

$$\frac{\partial P}{\partial x} = -i\omega\rho V_x \tag{2.33}$$

$$\frac{\partial P}{\partial y} = -i\omega\rho V_y \tag{2.34}$$

$$\frac{\partial P}{\partial z} = -i\omega\rho V_z \tag{2.35}$$

Equation 2.32 can be rearranged into

$$\frac{\partial V_z}{\partial z} = -\frac{i\omega}{K}P + \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} \tag{2.36}$$

Using 2.34 and 2.35 in 2.36, we have

$$\frac{\partial V_z}{\partial z} = -\frac{i\omega}{K}P + \frac{\partial}{\partial x}(\frac{1}{i\omega\rho}\frac{\partial P}{\partial x}) + \frac{\partial}{\partial y}(\frac{1}{i\omega\rho}\frac{\partial P}{\partial y}) \tag{2.37}$$

Equation 2.35 and equation 2.37 can be combined in matrix form:

$$\frac{\partial \mathbf{Q}}{\partial z} = \mathbf{AQ}, \tag{2.38}$$

where

$$\mathbf{Q} = \begin{pmatrix} -P \\ V_z \end{pmatrix} \tag{2.39}$$

and

$$\mathbf{A} = \begin{pmatrix} 0 & i\omega\rho \\ \frac{-1}{i\omega\rho}H_2 & 0 \end{pmatrix}, \tag{2.40}$$

where

$$H_2 = k^2 + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{\partial ln\rho}{\partial x}\frac{\partial}{\partial x} - \frac{\partial ln\rho}{\partial y}\frac{\partial}{\partial y}, \tag{2.41}$$

$$k^2 = \omega^2/c^2 \tag{2.42}$$

The propagation matrix **A** can be decomposed into the following form by eigenvalue decomposition:

$$\mathbf{A} = \mathbf{L} \wedge \mathbf{L}^{-1}, \tag{2.43}$$

where

$$\mathbf{L} = \begin{pmatrix} 1 & 1 \\ \frac{-1}{\omega\rho} H_1 & \frac{1}{\omega\rho} H_1 \end{pmatrix}, \tag{2.44}$$

$$\wedge = \begin{pmatrix} -iH_1 & 0 \\ 0 & iH_1 \end{pmatrix}, \tag{2.45}$$

and

$$\mathbf{L}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -\omega H_1^{-1}\rho \\ 1 & \omega H_1^{-1}\rho \end{pmatrix}, \tag{2.46}$$

where $H_1 = \sqrt{H_2}$. If the pressure field P can be decomposed into the downgoing wave P^+ and the updoing wave P^-, i.e., $P = P^+ + P^-$, then we can easily verify that

$$\mathbf{Q} = \mathbf{L}\mathbf{P}_l, \tag{2.47}$$

$$\mathbf{P}_l = \mathbf{L}^{-1}\mathbf{Q} \tag{2.48}$$

where

$$\mathbf{P}_l = \begin{pmatrix} -P^+ \\ -P^- \end{pmatrix}. \tag{2.49}$$

Equation 2.47 composes the total field from the upgoing and the dowgoing waves, and equation 2.48 decomposes the total wavefield into the downgoing wave and the upgoing wave. Therefore I call the two operators, $\mathbf{L}$ and $\mathbf{L}^{-1}$ given in equation 2.44 and 2.46, wave composition and wave decomposition operators respectively. Inserting equation 2.47 into 2.38 yields

$$\begin{aligned} \frac{\partial \mathbf{L}\mathbf{P}_l}{\partial z} &= \mathbf{A}\mathbf{L}\mathbf{P}_l \\ \frac{\partial \mathbf{L}}{\partial z}\mathbf{P}_l + \mathbf{L}\frac{\partial \mathbf{P}_l}{\partial z} &= \mathbf{L} \wedge \mathbf{L}^{-1}\mathbf{L}\mathbf{P}_l \\ \frac{\partial \mathbf{P}_l}{\partial z} &= (\wedge - \mathbf{L}^{-1}\frac{\partial \mathbf{L}}{\partial z})\mathbf{P}_l \end{aligned} \tag{2.50}$$

Using the expressions of $\mathbf{P}_l$, $\wedge$, $\mathbf{L}$ and $\mathbf{L}^{-1}$ in equation 2.50, we have two equations for upgoing and downgoing waves, respectively:

$$\frac{\partial P^+}{\partial z} = -iH_1 P^+ - \frac{1}{2} H_1^{-1} \Big[\rho \frac{\partial}{\partial z}(\frac{1}{\rho} H_1)\Big](P^+ - P^-) \tag{2.51}$$

$$\frac{\partial P^-}{\partial z} = +iH_1 P^- + \frac{1}{2} H_1^{-1} \Big[\rho \frac{\partial}{\partial z}(\frac{1}{\rho} H_1)\Big](P^+ - P^-) \tag{2.52}$$

It can be seen that if the term $\rho\frac{\partial}{\partial z}(\frac{1}{\rho}H_1)$ can be ignored, for example for a homogeneous medium, then we can completely decompose the wavefield into downgoing and upgoing waves. If only primary waves are considered, by ignoring the coupling between P^+ and P^-, then the decomposition is also valid, but only for sub-critical angles (Wapenaar and Berkhout, 1985). More general solution including critical angle events can be derived by adopting WKBJ representation, but the topic is beyond the scope of this thesis.

Equation 2.38 is called two-way wave equation, and the combination of 2.51 and 2.52 is called one-way wave equation. If the coupling is not ignored, two-way wave equation and one-way equation are equivalent. In practice, usually coupling of upgoing and downgoing is not considered for one-way wave equation. Therefore, two-way wave equation method is more accurate in modelling since multiples, transmission loss is automatically considered.

The solution of the two-way wave equation can be conveniently derived in the Fourier domain. In this domain, the two-way equation 2.38 can be formulated as:

$$\frac{\partial \tilde{\mathbf{Q}}}{\partial z} = \tilde{\mathbf{A}}\tilde{\mathbf{Q}}, \tag{2.53}$$

where

$$\tilde{\mathbf{A}} = \begin{pmatrix} 0 & i\omega\rho \\ -\tilde{H}_2/(i\omega\rho) & 0 \end{pmatrix}, \tag{2.54}$$

$$\tilde{\mathbf{Q}} = \begin{pmatrix} -\tilde{P} \\ \tilde{V}_z \end{pmatrix}, \tag{2.55}$$

$$\tilde{H}_2 = k_z^2 = k^2 - k_x^2 - k_y^2. \tag{2.56}$$

By eigenvalue decomposition, the matrix $\tilde{\mathbf{A}}$ can be decomposed into $\tilde{\mathbf{A}} = \tilde{\mathbf{L}}\tilde{\wedge}\tilde{\mathbf{L}}^{-1}$. The operators, $\tilde{\wedge}$, $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{L}}^{-1}$, are wavenumber-frequency domain counterparts of previously defined operators $\wedge$, $\mathbf{L}$ and $\mathbf{L}^{-1}$. These operators read

$$\tilde{\mathbf{L}} = \begin{pmatrix} 1 & 1 \\ \frac{-1}{\omega\rho\tilde{H}_1} & \frac{1}{\omega\rho\tilde{H}_1} \end{pmatrix}, \tag{2.57}$$

$$\tilde{\wedge} = \begin{pmatrix} -i\tilde{H}_1 & 0 \\ 0 & i\tilde{H}_1 \end{pmatrix}, \tag{2.58}$$

$$\tilde{L}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -\omega\tilde{H}_1^{-1}\rho \\ 1 & \omega\tilde{H}_1^{-1}\rho \end{pmatrix}, \tag{2.59}$$

where $\tilde{H}_1 = \sqrt{\tilde{H}_2} = \sqrt{k^2 - k_x^2 - k_y^2}$. For the wavenumber-frequency domain wave equation, Ursin (1983) proposed an extrapolation formula by assuming a homogeneous medium:

$$\tilde{\mathbf{Q}}(z) = \tilde{\mathbf{W}}(z, z_0)\tilde{\mathbf{Q}}(z_0), \tag{2.60}$$

where

$$\tilde{\mathbf{W}}(z, z_0) = e^{(\tilde{\mathbf{A}}\Delta z)}, \tag{2.61}$$

where z_0 is the depth of the surface, z is an arbitrary depth and $\Delta z = z - z_0$ is the thickness. Some manipulations of equation 2.61 yields (see Appendix A):

$$\tilde{\mathbf{W}}(z, z_0) = \begin{pmatrix} \tilde{W}_I(z, z_0) & \tilde{W}_{II}(z, z_0) \\ \tilde{W}_{III}(z, z_0) & \tilde{W}_{IV}(z, z_0) \end{pmatrix}, \tag{2.62}$$

where

$$\tilde{W}_I(z, z_0) = \cos(\tilde{H}_1 \Delta z), \tag{2.63}$$

$$\tilde{W}_{II}(z, z_0) = \frac{i\omega\rho}{\tilde{H}_1} \sin(\tilde{H}_1 \Delta z), \tag{2.64}$$

$$\tilde{W}_{III}(z, z_0) = \frac{1}{(\omega\rho)^2} \tilde{H}_2 \tilde{W}_{II}(z, z_0), \tag{2.65}$$

$$\tilde{W}_{IV}(z, z_0) = \tilde{W}_I(z, z_0). \tag{2.66}$$

The extrapolation method can be applied in an iterative way using:

$$\tilde{\mathbf{Q}}(z_i) = \tilde{\mathbf{W}}(z_i, z_{i-1})\tilde{\mathbf{Q}}(z_{i-1}) \tag{2.67}$$

Using above formula we can calculate the wavefield at depth $z = z_i$ from depth $z = z_{i-1}$. Similarly, we can derive a formula to calculate the wavefield at depth $z = z_{i-1}$ from $z = z_i$, which is used for modelling. The formula is

$$\tilde{\mathbf{Q}}(z_{i-1}) = \tilde{\mathbf{W}}(z_{i-1}, z_i)\tilde{\mathbf{Q}}(z_i), \tag{2.68}$$

where $\tilde{\mathbf{W}}(z_{i-1}, z_i)$ is in the same form as equation 2.62, but its components are computed by using $\Delta z = z_{i-1} - z(i)$. It is interesting to see that when $\tilde{H}_1 \to 0$, $\tilde{\mathbf{W}}(z, z_0)$ in equation 2.64 does not break down. The value of $\tilde{W}_{II}$ converges to $i\omega\rho\Delta z$, which allows us to use the formula to simulate head waves.

It is straightforward to show that the solution for one-way equation 2.51 and 2.52 are given by the following extrapolation formulas by assuming homogenous media:

$$\tilde{P}^{+}(z_i) = \tilde{P}^{+}(z_{i-1}) \cdot e^{-i\tilde{H}_1 \Delta z}, \tag{2.69}$$

$$\tilde{P}^{-}(z_i) = \tilde{P}^{-}(z_{i-1}) \cdot e^{+i\tilde{H}_1 \Delta z}. \tag{2.70}$$

Alternatively, equation 2.70 can be written as

$$\tilde{P}^{-}(z_{i-1}) = \tilde{P}^{-}(z_i) \cdot e^{-i\tilde{H}_1 \Delta z}. \tag{2.71}$$

These formulas can be used to propagate downgoing and upgoing waves independently in seismic simulation.

2.4.2 Two-way vs. one-way wave-equation modelling/migration

Two-way wave-equation modelling/migration

The recursive two-way representation of wave equation given by equation 2.68 can be utilized for seismic modelling. The wave extrapolation should not start from the surface since we don't have knowledge about the upgoing wave, the received reflection. Actually, if we know the upgoing wave at the surface, there is no need to do seismic modelling anymore. Therefore a rational scheme should start from the half space bellow the deepest layer. We can assume there is no upgoging wave in that area. The downgoing wave there is assumed to be an impulse response. The modelling procedure by two-way extrapolation operator is generalized as follows (Wapenaar and Berkhout, 1986b):

1. Prepare the total wavefield at the maximum depth $z = z_m$ by

   $$\tilde{\mathbf{Q}}(z_m) = \tilde{\mathbf{L}}\tilde{\mathbf{P}}(z_m), \tag{2.72}$$

   where $\tilde{\mathbf{P}}(z_m) = [\tilde{P}^{+}(z_m), \tilde{P}^{-}(z_m)]^T$. The downgoing wave $\tilde{P}^{+}(z_m)$ is set to an impulse response, and the upgoing wave $\tilde{P}^{-}(z_m)$ is set to zeros.

2. Use the recursive extrapolation equation 2.68 to propagate wavefield to the surface $z = z_0$.

3. Apply the following sequential equations:

$$\tilde{\mathbf{P}}(z_0) = [\tilde{P}^+(z_0), \tilde{P}^-(z_0)]^T = \tilde{\mathbf{L}}^{-1}(z_0)\tilde{\mathbf{Q}}(z_0), \tag{2.73}$$

$$\tilde{X}^{(0)}(z_0) = \tilde{P}^-(z_0)/\tilde{P}^+(z_0), \tag{2.74}$$

$$\tilde{X}(z_0) = [1 - \tilde{X}^{(0)}\tilde{R}^-(z_0)]^{-1}\tilde{X}^{(0)}(z_0), \tag{2.75}$$

   where the wavenumber-frequency domain upgoing reflectivity $\tilde{R}^-(z_0)$is given by

$$\tilde{R}^-(z_0) = \frac{\rho(z_0)\tilde{H}_1(z_1) - \rho(z_1)\tilde{H}_1(z_0)}{\rho(z_0)\tilde{H}_1(z_1) + \rho(z_1)\tilde{H}_1(z_0)} \tag{2.76}$$

4. Add signature of the source and receiver by

$$\tilde{P}^-_{CSG} = \tilde{S}(z_0)\tilde{X}(z_0)\tilde{D}(z_0), \tag{2.77}$$

   where $\tilde{S}(z_0)$ is the source signature, and $\tilde{D}(z_0)$ is the detector signature.

5. Repeat 1 ~ 4 for all wavenumbers and frequencies, then apply inverse Fourier transform to produce the common shot gather.

The extrapolation for two-way wave-equation migration starts from the surface. We assume that we can completely decompose the total wavefield at the surface into downgoing wave and upgoing wave. The former is set to an impulse response $\tilde{S}(z_0)^+$, and the latter is the received reflection $\tilde{P}_{CSG}(z_0)^-$. The two-way wave equation migration involves the following steps:

1. Calculate the total field $\tilde{\mathbf{Q}}(z_0)$ by the formula

$$\tilde{\mathbf{Q}}(z_0) = \tilde{\mathbf{L}}(z_0)\tilde{\mathbf{P}}(z_0) = \tilde{\mathbf{L}}(z_0)[\tilde{P}^+(z_0), \tilde{P}^-(z_0)]^T. \tag{2.78}$$

   For simplicity, we can use $\tilde{S}(z_0)^+$ as $\tilde{P}^+(z_0)$ and $\tilde{P}_{CSG}(z_0)^-$ as $\tilde{P}^-(z_0)$.

2. Use the recursive extrapolation equation 2.67 to calculate the total field at each depth $z = z_i$.

3. At this depth, we can calculate local downgoing and upgoing waves by

$$[\tilde{P}^+(z_i), \tilde{P}^-(z_i)]^T = \tilde{\mathbf{L}}(z_i)^{-1}\tilde{\mathbf{Q}}(z_i). \tag{2.79}$$

4. Apply the zero-offset imaging condition at this depth:

$$R(x,y,z_i) = \sum_{\omega} X(x,y,z_i,\omega), \tag{2.80}$$

where

$$X(x,y,z_i,\omega) = \frac{P^-(x,y,z_i,\omega)}{P^+(x,y,z_i,\omega)}, \tag{2.81}$$

where $P^-(x,y,z_i,\omega)$ and $P^+(x,y,z_i,\omega)$ are the space and temporal frequency domain wavefields converted from wavenumber-frequency domain wavefields $\tilde{P}^-(k_x,k_y,z_i,\omega)$ and $\tilde{P}^+(k_x,k_y,z_i,\omega)$, respectively, by inverse Fourier transform. Equation 2.81 may be unstable due to zero denominators. A more stable formula is as follows,

$$X(x,y,z_i,\omega) = \frac{P^{+*}(x,y,z_i,\omega) \cdot P^-(x,y,z_i,\omega)}{|P^+(x,y,z_i,\omega)|^2 + \sigma^2}, , \tag{2.82}$$

where $P^{+*}(x,y,z_i,\omega)$ is the adjoint of $P^+(x,y,z_i,\omega)$, and σ is a small constant to avoid infinite value of $X(x,y,z_i,\omega)$.

5. Repeat 2 $\sim$ 4 for all depths.

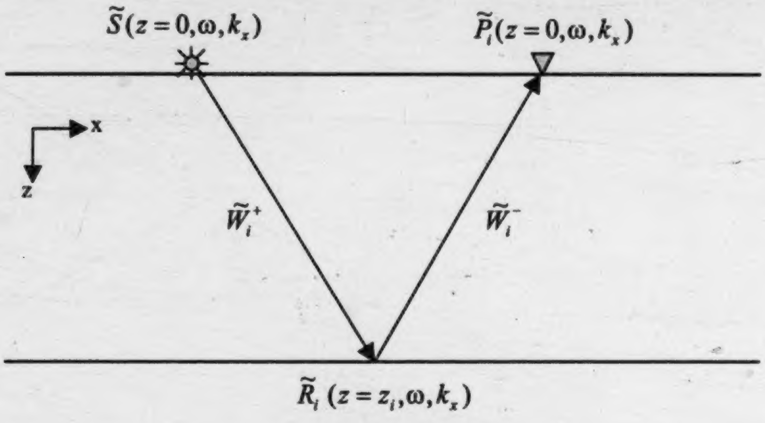
Although the above two-way wave-equation migration has potential to address multiples, transmission loss and converted waves, the method requires detail information about the velocity (Wapenaar and Berkhout, 1986b). Small velocity errors can make the algorithm unstable. For this reason, two-way wave-equation migration is not so robust as one-way wave equation migration described bellow.

One-way wave-equation modelling/migration

Berkhout (1984) proposed a scheme for one-way wave-equation modelling, which is portrayed in Figure 2.6. The procedure is given as follows:

1. The extrapolation starts from an impulse response at the source, symbolized with $\tilde{S}^+(z_0)$.

2. This wavefield is downward propagated by applying the recursive operator 2.69.


$\tilde{S}(z=0,\omega,k_x)$
$\tilde{P}_i(z=0,\omega,k_x)$
x
z
$\tilde{W}_i^+$
$\tilde{W}_i^-$
$\tilde{R}_i\ (z=z_i,\omega,k_x)$

$$\tilde{P}_i = \tilde{W}_i^- \tilde{R}_i \tilde{W}_i^+ \tilde{S}^+$$

$$\tilde{P} = \sum_i \tilde{P}_i$$

Figure 2.6: The scheme for one-way wave equation modelling. In the wavenumber-frequency domain, the contribution of each layer is considered a product of the source wavefield, downgoing operator, local reflectivity wavefield and upgoing propagator. The contributions of all depths are summed up to calculate the seismogram at the receiver. Finally, multi-dimensional inverse Fourier transform is applied to convert the data to the space and time domain.

3. When the wavefield reach a reflector (each depth is regarded as a reflector) at depth $z = z_i$, it is multiplied by the reflectivity wavefield $\tilde{R}_{z_i}$, which is calculated by

$$\tilde{R}(z_i) = \frac{\rho(z_i)\tilde{H}_1(z_{i-1}) - \rho(z_{i-1})\tilde{H}_1(z_i)}{\rho(z_i)\tilde{H}_1(z_{i-1}) + \rho(z_{i-1})\tilde{H}_1(z_i)} \tag{2.83}$$

4. Upward propagate the result wavefield of step 3 using the recursive operator 2.71 until it reaches the surface. The result wavefield is denoted by $\tilde{P}^-_{R_i}(z_0)$, which is the upgoing wave or the reflection excited by the reflector, R_i.

5. Repeat step 1 $\sim$ 4 for all reflectors (depths). Then sum up all reflected waves. The total reflection is calculated by

$$\tilde{P}^-(z_0) = \sum_{R_i} \tilde{P}^-_{R_i}(z_0) \tag{2.84}$$

6. Convert $\tilde{P}^-(z_0)$ to the space and time domain by inverse Fourier transform.

It can be seen that in the above scheme, we have ignored the transmission loss since we actually model the seismic response for each depth as if overlying layers do not exist, which is always violated in the real world. This is one disadvantage of this method.

One-way wave-equation migration also starts from the surface. It involves forward propagation of the source wavefield and backward propagation of the received reflection and application of imaging condition at each depth. The routine is generalized as bellow:

1. Calculate an impulse response at the receiver $\tilde{S}^+(z_0)$ in wavenumber-frequency domain. Convert the received reflection data to the same domain by Fourier transform. The result receiver wavefield is denoted as $\tilde{P}^-(z_0)$.

2. Use the recursive extrapolation formulas 2.69 and 2.70 to forward propagate the source wave and backward propagate the receiver wave, respectively. The result wavefields are denoted with $\tilde{P}^+(z_i)$ and $\tilde{P}^-(z_i)$

3. At each depth, the zero-offset imaging condition is applied:

$$R(x, y, z_i) = \sum_{\omega} X(x, y, z_i, \omega), \tag{2.85}$$

   where

$$X(x, y, z_i, \omega) = \frac{P^-(x, y, z_i, \omega)}{P^+(x, y, z_i, \omega)}, \tag{2.86}$$

   where $P^-(x, y, z_i, \omega)$ and $P^+(x, y, z_i, \omega)$ are the space and temporal frequency domain wavefields converted from wavenumber-frequency domain wavefields $\tilde{P}^-(k_x, k_y, z_i, \omega)$ and $\tilde{P}^+(k_x, k_y, z_i, \omega)$, respectively, by inverse Fourier transform. The more stable way to calculate $X(x, y, z_i, \omega)$ is as follows:

$$X(x, y, z_i, \omega) = \frac{P^{+*}(x, y, z_i, \omega) \cdot P^-(x, y, z_i, \omega)}{|P^+(x, y, z_i, \omega)|^2 + \sigma^2}, , \tag{2.87}$$

   where $P^{+*}(x, y, z_i, \omega)$ is the adjoint of $P^+(x, y, z_i, \omega)$, and σ is a small constant to avoid infinite value of $X(x, y, z_i, \omega)$.

4. Repeat 2 and 3 for all depth.

In this section, I have shown that the two-way and one-way wave equations are suitable for shot migration. However, as I have shown, both modelling and migration by the one-way approach are more explicitly related to the reflectivity. As a consequence, one-way wave-equation operators are more suitable for iterative inversion discussed in the next chapter, which aims at recovering a reflectivity model that can fit the seismic data. Since the described one-way method ignores multiples, multiples should be removed before migration/inversion.

For structural imaging, the described zero-offset imaging scheme is applied to all shot records. The contributions of all shots are added to produce an image cube (for 3-D) or a seismic profile (for 2-D). The summation can be formulated as bellow:

$$R(x,y,z) = \sum_{\mathbf{x}_s}\sum_{\omega} \frac{P^{-}(x,y,z,\omega,\mathbf{x}_s)P^{+*}(x,y,z,\omega,\mathbf{x}_s)}{P^{+}(x,y,z,\omega,\mathbf{x}_s)P^{+*}(x,y,z,\omega,\mathbf{x}_s)+\sigma^2} \tag{2.88}$$

The normalized imaging condition usually provides better solution than the conventional cross-correlation imaging condition (Claerbout, 1971) that does not consider the normalization of amplitude of downgoing waves. This is verified by a test of shot migration in Figure 2.7. The event especially at far offsets is better illuminated with the normalization. As shown later in this Chapter, the amplitude accuracy is also improved.

The imaging condition with amplitude normalization improves the image quality by introducing imaging deconvolution (Valenciano and Biondi, 2003). However, for iterative inversion, we should think the problem in a different way. By assuming the source wavefield as an impulse response, the one-way wave-equation modelling can be recast as:

$$\tilde{P} = \sum \tilde{W}_i^{-}\tilde{R}_i\tilde{W}^{+}, \tag{2.89}$$

where $\tilde{P}$ is the surface observation in the Fourier domain, $\tilde{W}_i$ is the upgoing propagation operator for the i th layer, $\tilde{R}_i$ is the reflector wavefield of the ith layer, and $\tilde{W}_i^{+}$ is the downgoing propagation operator for the ith layer. Therefore, the adjoint (migration) can be formulated as:

$$\tilde{R}_i = \tilde{W}_i^{-*}\tilde{P}\tilde{W}_i^{+*}, \tag{2.90}$$

where $\tilde{W}_i^{-*}$ is the adjoint of $\tilde{W}_i^{-}$, and $\tilde{W}_i^{+*}$ is the adjoint of $\tilde{W}_i^{-*}$.

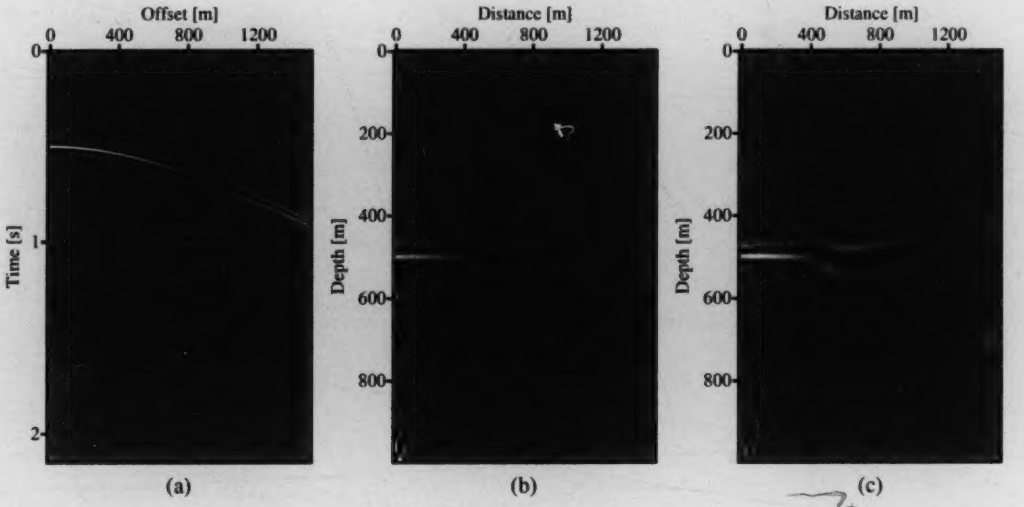


Figure 2.7: Migration of a single shot gather. (a) A shot gather with 61 offsets in a spacing of 25 m, sampled every $0.004s$. (b) Image obtained with cross-correlation imaging. (c) Image normalized by the amplitude of source wavefield.

2.5 Angle-dependent reflectivity

In the previous derivation of wave-equation modelling/adjoint, the concept of reflectivity is frequently used. It is important to understand the physical meaning of reflectivity before we use it for the forward and inverse problems. In an seismic experiment, usually we have many receivers at different lateral distance from the shot to detect the underground geological structure. The data acquisition system allows us to cover a survey range with rays of different angles. To understand this, we regard the seismic shot as a point source. Once the source is triggered, the energy is propagated toward all directions. One approach is to decompose the wavefield into plane waves and treat rays separately (Müler, 1971; Tygel and Hubral, 1984; Yilmaz and Taner, 1994). These rays impinge the reflector at different angles and rebound back to the surface. It is straight-forward to see that the received seismic response at receivers contains angle information. The more energy is reflected from the reflector, the stronger is the seismic response, which is rep-

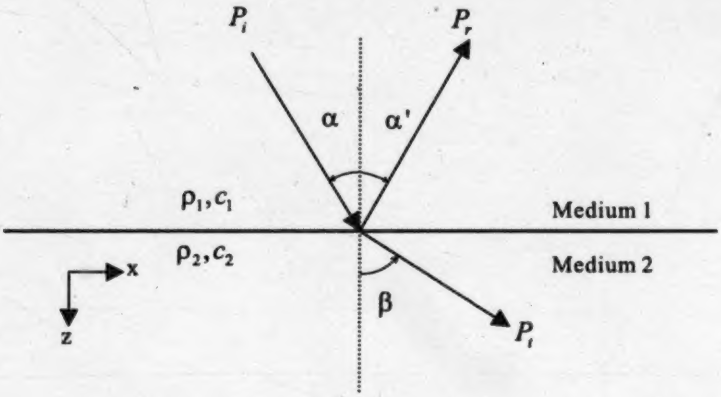


Figure 2.8: Diagram to derive angle-dependent reflectivity of a plane wave. P_i: the incident plane wave. P_r: the reflected plane wave. P_t: the transmitted plane wave. α: the incident angle. α': the reflection angle. β: the transmission angle.

resented by the amplitude of the seismic trace. By analyzing the boundary conditions of acoustic media at the reflector or the interface between two contacting layers, Berkhout (1984) proved that the amplitude can be angle-dependent.

Angle-dependent reflectivity

Figure 2.8 shows the geometry of one reflection event due to the incidence of a plane wave with a single frequency. The incident planar pressure wavefield $P_i(x, \omega)$ is represented by WKBJ approximation in space-frequency domain

$$P_i(x, \omega) = S_i(\omega) e^{-i\omega x / c_{x_1}}, \tag{2.91}$$

where S_i is the amplitude term, $e^{-i\omega x/c_x}$ is the phase term, and c_{x_1} is the apparent wave velocity in x direction. The apprarent velocity can be calculated by $c_{x_1} = c_1 / \sin \alpha$, where c_1 is the wave speed in the upper layer, and α is the incident angle. Similarly we can express the reflected and transmitted pressure wavefields by the following:

$$P_r(x, \omega) = S_r(\omega) e^{-i\omega x / c_{x_2}}, \tag{2.92}$$

$$P_t(x, \omega) = S_t(\omega) e^{-i\omega x / c_{x_3}}, \tag{2.93}$$

where S_r and S_t denote the amplitude of the reflected wave and transmitted wave respectively. The apparent velocities are related to the true velocity by:

$$\begin{aligned} c_{x_2} &= c_1/\sin\alpha', && (2.94) \\ c_{x_3} &= c_2/\sin\beta, && (2.95) \end{aligned}$$

where α' is the reflected angle, β is the transmission angle. c_2 is the true wave speed in the lower medium. Call two boundary conditions:

$$\begin{aligned} P_i + P_r &= P_t, && (2.96) \\ v_i \cos\alpha - v_r \cos\alpha' &= v_t \cos\beta. && (2.97) \end{aligned}$$

The first condition means that the pressure should be the same on both sides of the interface. The second condition uses the fact that particle velocity in normal direction should be continuous. Since the first condition holds for all frequency, the phase of equation 2.91, 2.92 and 2.93 should be consistent, which leads to Snell's law:

$$\frac{\sin\alpha}{c_1} = \frac{\sin\alpha'}{c_1} = \frac{\sin\beta}{c_2} = p, \tag{2.98}$$

where p is called ray parameter.

Using Snell's law in equation 2.96, we have

$$S_i + S_r = S_t. \tag{2.99}$$

The reflection coefficient or reflectivity is defined as $R = S_r/S_i$, and the transmission coefficient is defined as $T = S_t/S_i$. By using the definitions, equation 2.99 is reduced to;

$$1 + R = T. \tag{2.100}$$

To continue the derivation, we need a relationship between pressure, acoustic impedance and particle velocity (proved in Appendix B):

$$P = -Iv, \tag{2.101}$$

where $I = \rho c$, where ρ is the density and c is the wave speed. Using equation 2.101 in equation 2.97, we find

$$1 - R = \frac{T \cdot I_1 \cos\beta}{I_2 \cos\alpha}, \tag{2.102}$$

where $I_1 = \rho_1 c_1$ and $I_2 = \rho_2 c_2$. Combining equation 2.100 and equation 2.102, we have the formula for the angle dependent reflectivity:

$$R = \frac{I_2 \cos\alpha - I_1 \cos\beta}{I_2 \cos\alpha + I_1 \cos\beta} \tag{2.103}$$

The variable $\cos\beta$ can be replaced by

$$\begin{aligned} \cos\beta &= \sqrt{1 - \sin^2\beta} \\ &= \sqrt{1 - \frac{c_2^2}{c_1^2}\sin^2\alpha} \end{aligned} \tag{2.104}$$

Finally, the reflectivity is expressed as a function of the incident angle:

$$R(\alpha) = \frac{\rho_2 c_2 \cos\alpha - \rho_1 [c_1^2 - c_2^2 \sin^2\alpha]^{\frac{1}{2}}}{\rho_2 c_2 \cos\alpha + \rho_1 [c_1^2 - c_2^2 \sin^2\alpha]^{\frac{1}{2}}} \tag{2.105}$$

Snell's law tells us that the horizontal ray parameter is related to the incident angle. Therefore, alternatively we can express the angle dependent reflectivity in terms of the lateral ray parameter:

$$R(p) = \frac{\rho_2 c_2 \sqrt{1 - p^2 c_1^2} - \rho_1 \sqrt{c_1^2 - c_2^2 P^2 c_1^2}}{\rho_2 c_2 \sqrt{1 - p^2 c_1^2} + \rho_1 \sqrt{c_1^2 - c_2^2 P^2 c_1^2}}. \tag{2.106}$$

Equation 2.106 tells us that when more than four ray parameter dependant reflectivities are known (i.e. we have an over-determined system), we can invert the rock properties, ρ_1, ρ_2, c_1 and c_2. However, the inversion of rock properties is not the direct goal of my research. This thesis aims at recovering high-quality ray-parameter-dependent reflectivity model, $R(p)$ and puts forward the inversion of rock properties as an afterwards processing. For a discussion on how one can further invert for physical parameters see (Feng, 2004).

For an elastic earth model, we can similarly analyze the pressure and particle velocity of P-waves, S-waves and converted waves (Aki and Richards, 1980). The approximated solution for $\grave{P}\acute{P}$ wave is given in the introduction. The formula is obviously different from the acoustic case derived above. However, this does not influence the inversion method that I use since these formulas will be only used after the reflectivity model has been acquired by migration/inversion. As an early step of AVA analysis, the method adopted in this thesis inverts the compressional wave reflectivity by adopting the acoustic approximation of the elastic wave equation. Therefore Aki & Richard's formulation of AVA response is not involved.

local wavefield modelling and the adjoint

In the iterative inversion used in this thesis, a pair of conjugated operators should be developed to switch between data space and model space. As shown before, in one-way wave-equation modelling, downgoing wavefield from the source is multiplied by the local reflectivity wavefield to produce upgoing field in the wavenumber-frequency domain. In migration, the local reflectivity wavefield can be calculated by multiplying the upgoing wavefield by the adjoint of downgoing source wavefield. Therefore, the wavenumber domain local reflectivity wavefield at each depth can be regarded as the model that connects upgoing and downgoing wavefields. Since I am interested in the ray parameter dependent reflectivity model instead of the local wavefield, a pair of operators that relate the reflectivity model with the local wavefield are in order. One is used to synthesize local wavefield (wavenumber domain) from the reflectivity model (ray parameter domain) and the other one is to estimate reflectivity model from the local wavefield. These two operators can be designed by considering the Snell's law in the Fourier domain:

$$p = \frac{k}{\omega}, \tag{2.107}$$

To understand the problem, I denote the reflectivity model as $R(p, \omega, \mathbf{x}, z)$, and the local wavefield data as $\tilde{R}(k, \omega, \mathbf{x}, z)$, where $\mathbf{x}$ is the midpoint and z is the depth. In modelling of the local reflectivity wavefield, the amplitude in respect to a ray parameter is considered the same for all frequencies. Therefore, for each frequency, we can calculate the relative wavenumber k by equation 2.107 and set the amplitude of that wavenumber to the same value. Reversely, in the adjoint operation, we can calculate the reflectivity of one ray parameter by summing up the amplitude of all elements within the matrix $\tilde{R}(k, \omega, \mathbf{x}, z)$ that have same ratio of wavenumber and temporal frequency. In matrix form, the modelling procedure can be expressed as:

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{pmatrix} R(p,\mathbf{x},z) = \begin{pmatrix} \tilde{R}(k_1,\omega_1,\mathbf{x},z) \\ \tilde{R}(k_2,\omega_2,\mathbf{x},z) \\ \vdots \\ \tilde{R}(k_j,\omega_j,\mathbf{x},z) \\ \vdots \\ \tilde{R}(k_{N_{omega}},\omega_{N_\omega},\mathbf{x},z) \end{pmatrix},$$

where $k_j/\omega_j = p$. The adjoint operation is:

$$\bar{R}(p,\mathbf{x},z) = \begin{pmatrix} 1 & 1 & \cdots & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \tilde{R}(k_1,\omega_1,\mathbf{x},z) \\ \tilde{R}(k_2,\omega_2,\mathbf{x},z) \\ \vdots \\ \tilde{R}(k_j,\omega_j,\mathbf{x},z) \\ \vdots \\ \tilde{R}(k_{N_{omega}},\omega_{N_\omega},\mathbf{x},z) \end{pmatrix}, \tag{2.108}$$

where $\bar{R}(p,\mathbf{x},z)$ is the adjoint reflectivity model in the ray parameter domain. The adjoint operation is called radial trace transform (RTT) or slant stacking.

The described radial trace transform looks straight-forward, but it is not easy in application. The first problem is the mismatch between regularly sampled frequency, wavenumber and ray parameter. For example, we have a common midpoint gather with 40 offsets in a spacing of 25 m, 400 samples per trace and a sampling rate of $4ms$. The discrete frequency is $(0, 3.927, 7.854, ...2\pi \cdot 125)$. The wavenumber is $(0, 0.0063, 0.0126, ..., 2\pi \cdot 0.02)$. As an example, I set the ray parameter to $(0, 0.002, 0.004, ...)$. The relative wavenumber vector of the given ray parameter for the first frequency 3.927 is $(0, 0.0079, 0.0158, ...)$. It is clear that most of elements are not exactly equal to any elements of the defined wavenumber vector. The solution of this problem is interpolation (Claerbout, 1985). The amplitude of one wavenumber is calculated by linearly interpolating the amplitude of neighboring wavenumber. For example, the wavenumber 0.0079 is located between 0.0063 and 0.0126. Therefore, its amplitude is calculated by:

$$R(0.079) = \frac{0.0126 - 0.0079}{0.0126 - 0.0063} R(0.0063) + \frac{0.0079 - 0.0063}{0.0126 - 0.0063} R(0.0126). \tag{2.109}$$

The second problem is the bandlimited nature of wavenumber spectra. According to Nyquist law, the maximum offset wavenumber is $1/(2dh)$. This means that when the product of ray parameter and frequency is larger than the wavenumber up-limit, we can not calculate the amplitude of this ray parameter at this frequency. The scenario usually happens at high angles or large ray parameters, which leads to inaccurate amplitude of high angles. This is a very common problem for all kinds of linear imaging problems. In Chapter 3, I will show that the accuracy at high angles can be improved by introducing non-linear regularization.

A ray parameter sampling rule for aliasing-free image gathers

Since the reflectivity model of each ray parameter gives an image of the earth structure, we can define a concept, image gather, for the gather of ray parameter dependent reflectivity. The gather for one midpoint is called Common Image Gather (CIG). To avoid aliasing of CIGs, the ray parameter should be adequately sampled to make full use of available information from the given data. To see this, I write the Snell's Law in the Fourier domain as bellow,

$$p_h = \frac{k_h}{\omega} \tag{2.110}$$

where p_h is the ray parameter in offset direction, k_h is the offset wavenumber and ω is the temporal frequency. In the real world, we have discrete signals. Both wavenumber k_h and ray parameter p_h are usually sampled in regular spacings. At a given frequency ω, they should obey

$$dp_h = \frac{dk_h}{\omega} \tag{2.111}$$

where dp_h is the ray parameter spacing, and dk_h is the wavenumber spacing. Equation 2.111 gives a good sampling standard to avoid aliasing for linear inversion with ray parameter and wavenumber as the input and output. The good fitting between two sides may also facilitate the convergence of iterative inversion. One straightforward application is to apply the sampling rule to time-invariant transforms like radial trace stacking (Kostov, 1990). If k_h is known and we want to solve p_h, ray parameter spacing should obey:

$$dp_h \leq \frac{dk_h}{\omega} = \frac{2\pi/(N_h \cdot dh)}{2\pi \cdot f} = \frac{1}{N_h \cdot dh \cdot f} \tag{2.112}$$

where N_h is the offset number, dh is the offset spacing and f is the frequency. This can be extended to band-limited signals:

$$dp_h \leq \frac{1}{N_h \cdot dh \cdot f_{max}} \tag{2.113}$$

where f_{max} is the maximum frequency. It is clear that when the above inequality is fulfilled, ray parameters for all frequencies f are well sampled since $f \leq f_{max}$.

To verify this ray parameter sampling rule, I examine the influence of ray parameter sampling on linear slant stacking Radon inversion in the Fourier domain. The data have 120 offsets with 5 m spacing. Each seismic trace has 500 samples, and the sampling rate is $4ms$. The maximum frequency is 80 Hz. Some hyperbolic events are prepared to simulate a common midpoint gather. Using the sampling rule, the largest ray parameter spacing is $1/(120 \times 5 \times 80) = 0.000021 s/m = 21 \mu s/m$. A larger interval of the ray parameter will introduce aliasing in the solution. Figure 2.9a is the original input data. Figure 2.9 b and c display the inverted radon panels with a good sampling and a bad sampling of the ray parameter, respectively. It's clear that with good ray parameter sampling, the model is better focused.

Figure 2.9d and e show the predicted data by applying the forward operator to the inverted radon panels. The result of the well sampled solution successfully recover the data and the residual (see Figure 2.9f) is ignorable compared with the original data. On the other hand, under-sampled solution introduces many artifacts in the modelled data. The data fitting is not acceptable with a significant amount of coherent events present in the residual (see Figure 2.9g).

The sampling criterion can be used for AVP (Amplitude Versus ray Parameter) wave-equation migration. In AVP migration, the imaging procedure utilizes the radial trace transform at each depth, and the spacing of wavenumber does not change for all depths, which means once the surface ray parameters are well sampled, so are those of the subvolume.

2.6 Angle-domain imaging condition

With the concept of local reflectivity field modelling (the adjoint of RTT) and the adjoint (RTT), we can now better understand the procedure of shot profile modelling and

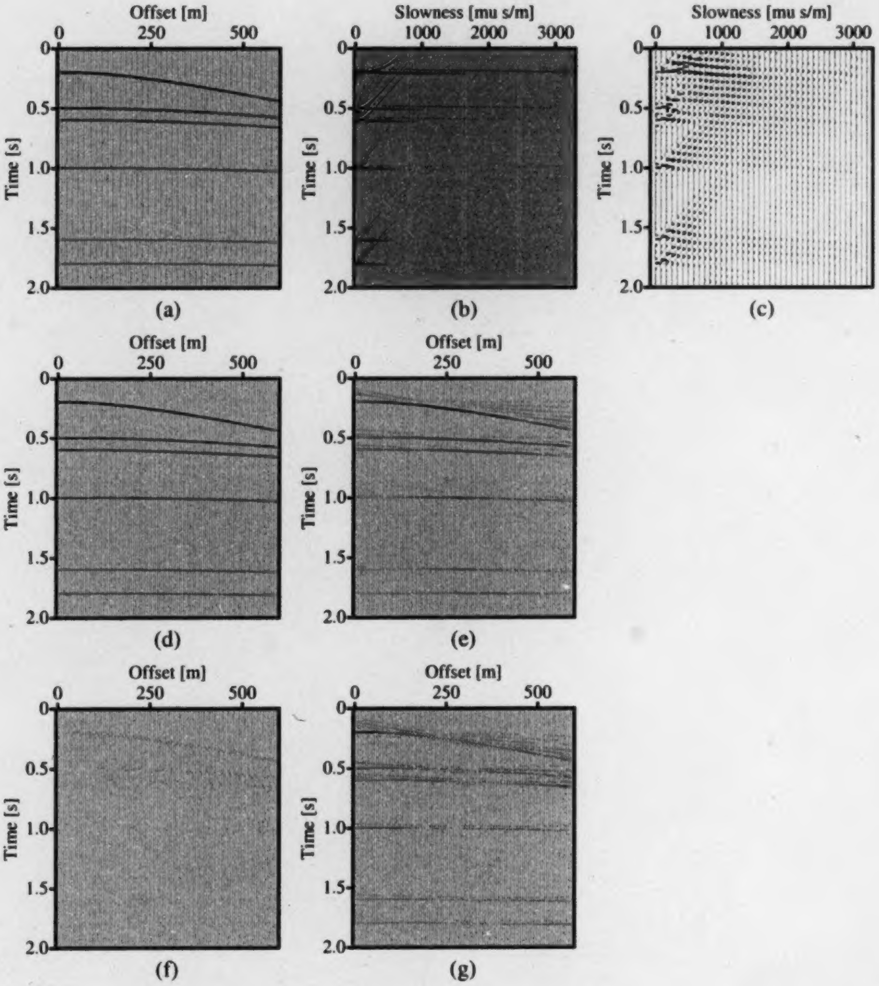


Figure 2.9: A common midpoint gather, the inverted radon panels, reconstructed data and the residuals. (a) The synthetic data. (b) Radon panel with good sampling of ray parameter ($dp = 21\mu s/m$). (c) Radon panel with bad sampling of ray parameter ($dp = 84\mu s/m$) (d) Reconstructed data with (b). (e) Reconstructed data with (c). (f) Residual (a)-(d). (g) Residual (a)-(e).

migration. In modelling, we regard each depth as an perturbation source. First we downward continue the impulse response from the source to each depth. At each depth, we invoke the the adjoint of RTT to produce a local reflectivity wavefield and multiply it with the downward continued source wavefield. Then the result is upward continued to the surface. At surface, the contributions from all depths are added to calculate the final shot gather. In migration, we forward propagate the source wavefield and backward propagate the receiver wavefield to each depth. The local wavefield is calculated by multiplying the downward continued receiver wavefield with the adjoint of the downward continued source wavefield. Then common image gathers can be produced by RTT:

$$R(p_x, p_y, z_i) = \sideset{}{'}\sum_{\omega} \tilde{X}[k_x, k_y, z_i, \omega], \tag{2.114}$$

where

$$\tilde{X}[k_x, k_y, z_i, \omega] = \tilde{P}^{-}(k_x, k_y, z_i, \omega) \cdot \tilde{P}^{+*}(k_x, k_y, z_i, \omega) \tag{2.115}$$

The prime in equation 2.114 means slant stacking or RTT, which has been described in the previous sub-section.

The imaging condition 2.114 contains angle-domain information in terms of ray parameters. However, it is only suitable for lateral invariant media since different lateral positions should have different $p-z$ panels due to velocity variation. The formula cannot tell the position of the resulting imaging gathers. Alternatively, Rickett and Sava (2002) provided a new scheme to produce angle-domain image gathers for shot-profile migration, which involves sorting offsets for each midpoint and applying imaging condition separately. The scheme can be generalized as the following steps:

1. Extrapolate the source field and receiver field to each depth. The wavefields are denoted as $P^{+}(x, y, z, \omega)$ and $P^{-}(x, y, z, \omega)$, respectively.

2. Apply the cross-correlation image condition using downgoing waves and upgoing waves with variant displacement. The formula of the imaging condition reads

$$R(\mathbf{x}, \mathbf{h}, z) = \sum_{s}\sum_{\omega} P^{-}(\mathbf{x} - \mathbf{h}, z, \omega, s) P^{+*}(\mathbf{x} + \mathbf{h}, z, \omega, s). \tag{2.116}$$

In this way, we can calculate multi-offset versus depth ($\mathbf{h} - z$) panels for all midpoints. The result is AVO gathers in depth domain, different from traditional AVO gathers in time domain by using simpler operators like NMO or DMO..

3. Convert AVO image gathers from the space domain ($R(\mathbf{h}, z)$) to the Fourier domain ($\tilde{R}(\mathbf{k}_h, kz)$).

4. Convert AVO gathers to AVA gathers using the relationship between wavenumber and wave incident angle (Sava and Fomel, 2000):

$$\tan\gamma = \frac{|\mathbf{k}_h|}{k_z}, \tag{2.117}$$

   where γ is the incident angle. The conversion is applied in two steps. The first step is to use the above equation to calculate amplitude of different angles for each vertical wavenumber, i.e., change the image from (k_z, k_h) domain to (k_z, γ) domain for each midpoint. The second step is to apply inverse Fourier transform to compute the depth variant angle image gather, i.e., convert the image from (k_z, γ) domain to (z, γ) domain.

The above approach is called "image-space" method (Sava and Fomel, 2000; Fomel, 2004) since the offset to angle conversion is applied after migration. Another method converts the image in the data space during migration by using the previously discussed imaging condition, radial trace transform (Prucha et al., 1999). The method is applied in the following steps:

1. Extrapolate the source field and receiver field to each depth. The wavefields are denoted as $P^{+}(x, y, z, \omega)$ and $P^{-}(x, y, z, \omega)$, respectively.

2. Apply the cross-correlation image condition using downgoing waves and upgoing waves with variant displacement for each frequency. The formula of the imaging condition reads

$$R(\mathbf{x}, \mathbf{h}, \omega, z) = \sum_{s} P^{-}(\mathbf{x} - \mathbf{h}, z, \omega, s) P^{+*}(\mathbf{x} + \mathbf{h}, z, \omega, s). \tag{2.118}$$

3. Convert offset image gathers to offset wavenumber image gathers by 1-D Fourier transform along the offset direction, i.e., change $R(\mathbf{x}, \mathbf{h}, z, \omega)$ to $\tilde{R}(\mathbf{x}, k_h, \omega, z)$.

4. Produce ray parameter image gathers by RTT. The result is denoted as $R(\mathbf{x}, \mathbf{p}_h, z)$

5. Convert ray parameter image gathers to angle image gather by the following relationship (Rickett and Sava, 2002):

$$\mathbf{p}_h = \begin{pmatrix} p_{h_x} \\ p_{h_y} \end{pmatrix} = \frac{2\sin\gamma\cos\theta}{v(\mathbf{x}, z)} \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}, \tag{2.119}$$

   where p_{h_x} is the inline offset wavenumber, p_{h_y} is the crossline offset wavenumber, γ is the incident angle, θ is the structural dip, $v(\mathbf{x}, z)$ is the wave speed at the reflection point and ϕ is the azimuth angle of the plane wave. The azimuth angle ϕ is defined as the angle between inline direction and the maximum down-dip direction at the reflector point.

In 3-D common-azimuth migration described in section 2.8, we only need to worry about the inline offset ray parameter p_{h_x}. If we ignore the azimuth angle ϕ, then the formula is simplified into

$$p_{h_x} = \frac{2\sin\gamma\cos\theta}{v(\mathbf{x}, z)}. \tag{2.120}$$

Please note that for linear inversion, the adjoint operator (migration) only requires step 1 ~ 4. When a good model of ray parameter dependent image gathers has been acquired by the inversion, we can convert it to the angle dependent model by equation 2.119 or equation 2.120.

At first glance, the model-space method of producing angle image gathers is more attractive since no structural dip information is needed. But the method is more sensitive to velocity errors (Sava and Fomel, 2003).

Fomel (2004) proposed a more general form of the relationship between wavenumbers and the incident angle:

$$\begin{aligned} &(k_{m_x}^2 + k_{m_y}^2)\frac{\sin^2\gamma}{v^2} + (k_{h_x}^2 + k_{h_y}^2)\frac{\cos^2\theta}{v^2} = \\ &\frac{1}{4\omega^2}(k_{m_x}k_{h_y} - k_{m_y}k_{h_x})^2 + 4\omega^2\frac{\cos^2\theta}{v^2}\frac{\sin^2\theta}{v^2}, \end{aligned} \tag{2.121}$$

where k_{m_x} and k_{m_y} are inline and crossline midpoint wavenumbers, and k_{h_x} and k_{h_y} are inline and crossline offset wavenumbers. For a 2-D model, the above dispersion equation can be simplified by setting k_{m_y} and k_h to zeros:

$$k_h = \frac{2\omega \sin\gamma}{v}\sqrt{1 - \frac{4k_m^2 v^2}{\omega^2 \cos^2\gamma}}. \tag{2.122}$$

Interestingly, it can be proved (Fomel, 2004) that common-azimuth migration can use the same formula as above to convert ray parameter to incident angle. The formula is attractive since no knowledge about the structural dip is required. However it is applied in the midpoint wavenumber domain, which implicitly assumes constant velocity in the lateral direction. For this reason, it is difficult to link the parameter v to local velocity at the imaging point.

2.7 DSR wave-equation modelling/migration

2.7.1 Theory of DSR modelling/migration

Seismic data obtained via a multi-sources multi-receivers experiment can be sorted and processed together to image the earth interior (Claerbout, 1985). The data at the surface is considered as a function of the source and the receiver position, denoted as $P(z = 0, t, \mathbf{s}, \mathbf{g})$, where P is the pressure, z is the depth, t is the time, $\mathbf{s}$ is the source vector, and $\mathbf{g}$ is the receiver vector. When the recorded data are converted to the Fourier domain, extrapolation formulas can be adopted to link wavefields at different depths. The idea is used to develop the theory of Double-Square-Root (DSR) wave-equation migration (Claerbout, 1985). For simplicity, I show bellow the derivation of this method for 2-D case. The 3-D formula (Biondi, 2003a) is given after the derivation. Consider a plane wave propagating at an incident angle of θ (see Figure 2.10). It is clear that the horizontal apparent velocity of the wave is $v_x = v/\sin\theta$ and the vertical apparent velocity is $v_z = v/\cos\theta$, which leads to two important relationships:

$$\frac{1}{v_x} = \frac{\partial t}{\partial x} = \frac{\sin\theta}{v} = p \tag{2.123}$$

$$\frac{1}{v_z} = \frac{\partial t}{\partial z} = \frac{\cos\theta}{v} = \sqrt{\frac{1}{v^2} - (\frac{\partial t}{\partial x})^2} \tag{2.124}$$

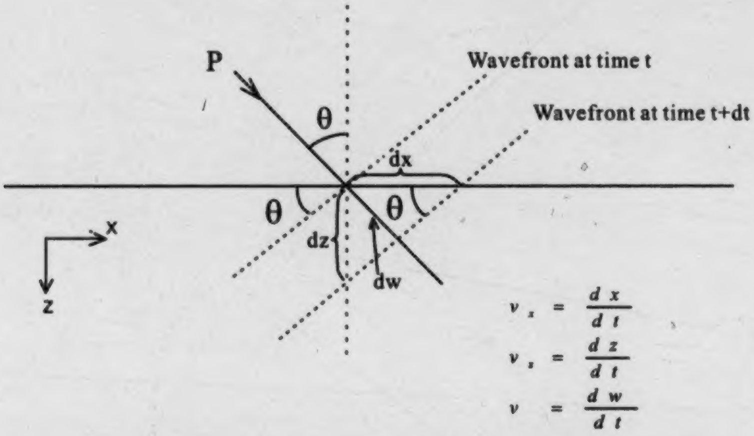


Figure 2.10: Derivation of apparent velocities. A plane wave impinges the interface with an angle θ. It is clear that the angle between the wavefront and the horizontal direction is identical to the incident angle. Simple geometry leads to the relationship between the horizontal and vertical apparent velocities (v_x and v_z) and the true wave speed v: $v_x = v/\sin\theta, v_z = v/\cos\theta$.

If both source and receiver depths are considered, then the vertical slowness should be

$$\begin{aligned}\frac{\partial t}{\partial z} &= \frac{\partial t}{\partial z_g} + \frac{\partial t}{\partial z_s} \\ &= -\Big[\sqrt{\frac{1}{v^2} - (\frac{\partial t}{\partial g})^2} + \sqrt{\frac{1}{v^2} - (\frac{\partial t}{\partial s})^2}\Big], \end{aligned} \tag{2.125}$$

where the second identity is derived by using equation 2.124 for the source and receiver respectively. The time-shifting partial-differential equation for the upgoing wave is

$$\frac{\partial P}{\partial z} = -\frac{\partial t}{\partial z}\frac{\partial P}{\partial t}. \tag{2.126}$$

Inserting equation 2.125 into equation 2.126 yields

$$\frac{\partial P}{\partial z} = \Big[\sqrt{\frac{1}{v^2} - (\frac{\partial t}{\partial g})^2} + \sqrt{\frac{1}{v^2} - (\frac{\partial t}{\partial s})^2}\Big]\frac{\partial P}{\partial t}. \tag{2.127}$$

One trick to solve partial derivative equations in the Fourier domain is to replace the derivative sign with the product of relative wavenumber and the complex number i. For example, $1/\partial g$ is replaced with ik_g, and $1/\partial s$ is replaced with ik_s. The temporal wavenumber is a little different. The time differential operator $1/\partial t$ is replaced with $-i\omega$. Here I use the Claerbout's sign convention of Fourier transform for time and temporal frequency domain. With this sign convention, Fourier transform and inverse Fourier transform are listed bellow

$$\tilde{P}(\omega) = \int P(t) e^{j\omega t} dt \tag{2.128}$$

$$P(t) = \frac{1}{2\pi} \int \tilde{P}(\omega) e^{-j\omega t} d\omega. \tag{2.129}$$

On the other hand, spatial and wavenumber domain Fourier transform and inverse Fourier transform are defined in the usual way with the familiar sign convention. Take the receiver wavefield as an example:

$$\tilde{P}(k_g) = \int P(g) e^{-jk_g g} dg \tag{2.130}$$

$$P(g) = \frac{1}{2\pi} \int \tilde{P}(k_g) e^{jk_g g} dk_g \tag{2.131}$$

By using the wavenumber replacement trick, equation 2.127 is reduced to

$$\begin{aligned} \frac{\partial \tilde{P}}{\partial z} &= -j\omega \Big[\sqrt{\frac{1}{v^2} - (\frac{k_g}{\omega})^2} + \sqrt{\frac{1}{v^2} - (\frac{k_s}{\omega})^2} \Big] \tilde{P} \\ &= -j \Big[\sqrt{\frac{\omega^2}{v^2} - k_g^2} + \sqrt{\frac{\omega^2}{v^2} - k_s^2} \Big] \tilde{P} \end{aligned} \tag{2.132}$$

The equation can be converted to midpoint and offset wavenumber domain by the following relationships:

$$k_g = \frac{1}{2}(k_m + k_h) \tag{2.133}$$

$$k_s = \frac{1}{2}(k_m - k_h), \tag{2.134}$$

where k_m is the midpoint wavenumber and k_h is offset wavenumber.The resulting equation is

$$\begin{aligned} \frac{\partial \tilde{P}}{\partial z} &= -j \Big[\sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_m + k_h)^2} + \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_m - k_h)^2} \Big] \tilde{P} \\ &= -jk_z \tilde{P}, \end{aligned} \tag{2.135}$$

where k_z denotes the two square roots. It can also be called the vertical wave number.

The recursive numerical solution of the above equation is

$$\tilde{P}(z_i) = \tilde{P}(z_{i-1})e^{-ik_z}\Delta z \tag{2.136}$$

For 3-D case, the vertical wave number k_z is formulated as (Biondi, 2003a)

$$\begin{aligned} k_z &= \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}\left[(k_{m_x} + k_{h_x})^2 + (k_{m_y} + k_{h_y})^2\right]} \\ &+ \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}\left[(k_{m_x} - k_{h_x})^2 + (k_{m_y} - k_{h_y})^2\right]}\,, \end{aligned} \tag{2.137}$$

where k_{m_x} and k_{m_y} are inline and crossline midpoint wavenumbers respectively. k_{h_x} and k_{h_y} are inline and crossline offset wavenumbers. The offset here is half offset, which means the distance between the midpoint and the source or receiver.

For lateral invariant media, we can further simplify the solution into

$$\tilde{P}(z_i) = \tilde{P}(z_0)e^{-j\sum_{l=1}^{i} k_{z_l}\Delta z} \tag{2.138}$$

This is exactly the same as the local reflectivity wavefield that has been introduced in the section of shot-profile migration. I prove the statement as follows.

When the media is laterally invariant, the midpoint wavenumbers disappear according to the theory of discrete Fourier transform, i.e., $k_{m_x} = 0$ and $k_{m_y} = 0$. Consequently, the above vertical number k_{zl} is reduced to

$$k_{zl} = \sqrt{\frac{\omega^2}{v_l^2} - (k_{h'_x}^2 + k_{h'_y}^2)} + \sqrt{\frac{\omega^2}{v_l^2} - (k_{h'_x}^2 + k_{h'_y}^2)}, \tag{2.139}$$

where

$$k_{h'_x} = 0.5k_{h_x}, \tag{2.140}$$

$$k_{h'_y} = 0.5k_{h_y}. \tag{2.141}$$

It is easy to show that if $h'_x = 2h_x$ and $h'_y = 2h_x$, then the wavenumber condition 2.140 and 2.141 are fulfilled. For example, we have two data sets, both of which have nh offsets. The spacings of them are dh and 0.5dh, respectively. The wavenumber of one dataset should be $(0, 2\pi/(nh \cdot dh), 4\pi/(nh \cdot dh), ...)$. The other one is $(0, 4\pi/(nh \cdot dh), 8\pi/(nh \cdot dh), ...)$.

In Claerbout's sign convention, the recursive extrapolation equations for downgoing wave and upgoing wave are

$$\begin{aligned} \tilde{P}^{+}(z_i) &= \tilde{P}^{+}(z_{i-1}) \cdot e^{j\tilde{H}_1 \Delta z}, & (2.142) \\ \tilde{P}^{-}(z_i) &= \tilde{P}^{-}(z_{i-1}) \cdot e^{-j\tilde{H}_1 \Delta z}, & (2.143) \end{aligned}$$

As shown before, the local reflectivity wavefield is calculated by

$$\tilde{R}(z_i) = \tilde{P}^{-}(z_i) \cdot \tilde{P}^{+*}(z_i). \tag{2.144}$$

For lateral invariant media, the equation can be reformulated as

$$\begin{aligned} \tilde{R}(z_i) &= \tilde{P}^{-}(z_0) \cdot e^{-j\sum_{l=1}^{i} \tilde{H}_{1l} \Delta z} \,.\, \tilde{P}^{+}(z_0) \cdot e^{-j\sum_{l=1}^{i} \tilde{H}_{1l} \Delta z} \\ &= \tilde{P}^{-}(z_0) \cdot e^{-j(\sum_{l=1}^{i} \tilde{H}_{1l} \Delta z + \sum_{l=1}^{i} \tilde{H}_{1l} \Delta z)}. \end{aligned} \tag{2.145}$$

The second identity is derived by assuming the source field is an impulse response so that $\tilde{P}^{+}(z_0) = 1$. Recall that the formula for $\tilde{H}_{1l}$ is

$$\tilde{H}_{1l} = \sqrt{\frac{\omega}{v_l}^2 - (k_x^2 + k_y^2)}. \tag{2.146}$$

The image grid spacing of a shot gather is two times of that of the common midpoint gather. Therefore, $k_x = 0.5k_{h_x} = k_{h'_x}$ and $k_y = 0.5k_{h_y} = k_{h'_y}$. Finally, the local reflectivity wavefield can be calculated by

$$\tilde{z}_i = \tilde{P}^{-}(z_0) e^{-j\sum_{l=1}^{i} k_{zl} \Delta z}, \tag{2.147}$$

where

$$k_{zl} = \sqrt{\frac{\omega^2}{v_l^2} - (k_{h'_x}^2 + k_{h'_y}^2)} + \sqrt{\frac{\omega^2}{v_l^2} - (k_{h'_x}^2 + k_{h'_y}^2)}. \tag{2.148}$$

Comparing the formulas of DSR migration and shot-profile migration, we can see that when velocity does not change in the lateral direction, two methods should provide identical result. Please note that a proof in the space and temporal frequency domain is given by Biondi (2003b)

Equation 2.136, 2.137 and the previously discussed theory of local reflectivity wavefield provide the foundation for DSR wave-equation modelling and migration. The modelling is generalized as the following steps:

1. At each depth, apply the adjoint of Radial Trace Transform (RTT) to calculate the local reflectivity wavefield $\tilde{P}(z_i, \omega, \mathbf{k}_h, \mathbf{m})$. Then apply fourier transform to convert the local wavefield to midpoint wavenumber domain, i.e., change from $\tilde{P}(z_i, \omega, \mathbf{k}_h, \mathbf{m})$ to $\tilde{P}(z_i, \omega, \mathbf{k}_h, \mathbf{k}_m)$

2. Recursively use the phase-shift operator to propagate the wavefield to the surface ($z = z_0$). The formula of recursive DSR modelling is acquired by rearranging equation 2.136:

$$\tilde{P}(z_{i-1}) = \tilde{P}(z_i)e^{i\mathbf{k}_z\Delta z}. \tag{2.149}$$

3. Repeat 1 and 2 for all depths. When all wavefields have been propagated to the surface, they are summed up and converted to the space and time domain.

The adjoint operator (migration) involves applying the phase-shift wave propagator with negative sign. It is done by the following steps:

1. Convert the seismogram received at the surface to the wavenumber and frequency domain. The resulting data are denoted by $\tilde{P}(z = 0, \omega, \mathbf{k}_m, \mathbf{k}_h)$.

2. Recursively apply the downward continuation operator to each depth. The formula is given by equation 2.136.

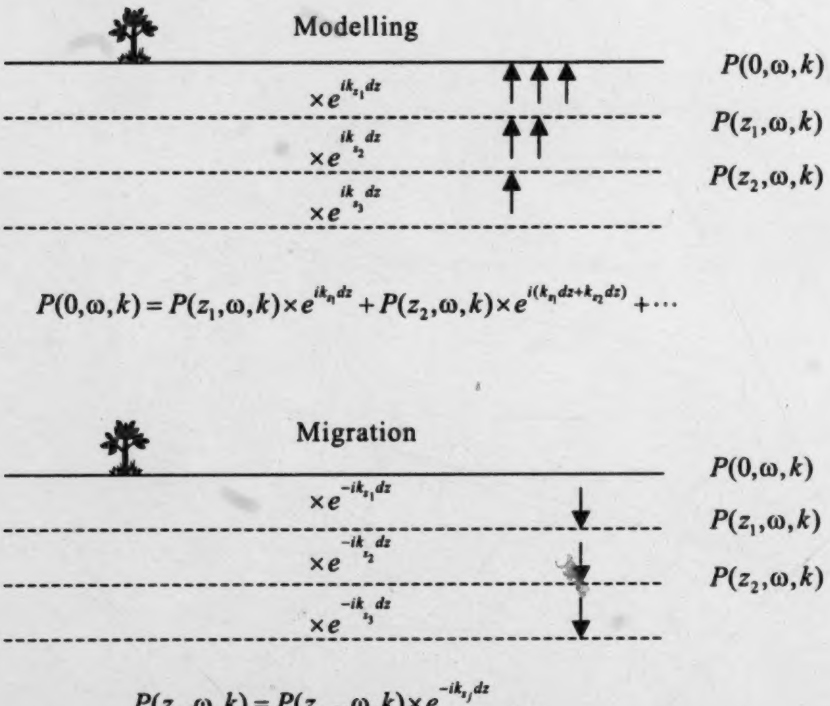
3. At each depth, RTT is used to compute common image gathers.

Figure 2.11 portrays these two conjugated procedures.

The general form of DSR operator can be changed to operators of 3-D zero-offset migration, 2-D prestack migration, 2-D poststack migration and 3-D common-azimuth migration by setting some wavenumbers to zeros or using some approximation (see Table 2.1). The listed common-azimuth migration method involves a stationary-phase approximation of the crossline offset wavenumber, which provides an efficient solution to the 3-D imaging problem. More detail about this operator is given in the section "common-azimuth approximation".

Migration type	Wavenumber values
3-D zero-offset (poststack) migration	$k_{h_x} = 0, k_{h_y} = 0$
2-D prestack migration	$k_{m_y} = 0, k_{h_y} = 0$
2-D poststack migration	$k_{m_y} = 0, k_{h_x} = 0, k_{h_y} = 0$
3-D prestack common-azimuth migration	Replace k_{h_y} with $\hat{k}_{h_y}$

Table 2.1: Some migration operators and their wavenumber values


Modelling
$P(0,\omega,k)$
$\times e^{ik_{z_1}dz}$
$P(z_1,\omega,k)$
$\times e^{ik_{z_2}dz}$
$P(z_2,\omega,k)$
$\times e^{ik_{z_3}dz}$
Migration
$P(0,\omega,k)$
$\times e^{-ik_{z_1}dz}$
$P(z_1,\omega,k)$
$\times e^{-ik_{z_2}dz}$
$P(z_2,\omega,k)$
$\times e^{-ik_{z_3}dz}$

$$P(0,\omega,k) = P(z_1,\omega,k)\times e^{ik_{z_1}dz} + P(z_2,\omega,k)\times e^{i(k_{z_1}dz+k_{z_2}dz)} + \cdots$$

$$P(z_j,\omega,k) = P(z_{j-1},\omega,k)\times e^{-ik_{z_j}dz}$$

Figure 2.11: Schemes of Double-Square-Root (DSR) wave equation modeling and migration. In modeling, the surface wavecan be considered as the summation of propagated waves from all depths. The steps of propagations vary with the depth. In migration (the adjoint of modeling), the wave of each depth is calculated by downward continuing the surface wave recursively.

2.7.2 A toy example of AVP migration

To show the similarity between shot-profile migration and DSR migration, I prepared a 2-D synthetic data set. The model consists of one layer and one half-space. Their parameters are listed in Table 2.2.

Velocity (m/s)	Density (g/cm^3)	Thickness (m)
2000	2.25	500
2350	1.6	Half-space

Table 2.2: Model parameters for the 2-D synthetic data.

A shot record with 61 offsets was produced by a reflectivity modelling method for acoustic media. The shot gather is displayed in Figure 2.7a. The shot gather was copied ten times to simulate a 2-D survey inline. I compare the ray parameter dependent (AVP) image gathers from shot-profile migration and DSR migration (see Figure 2.12). They are almost identical in shape.

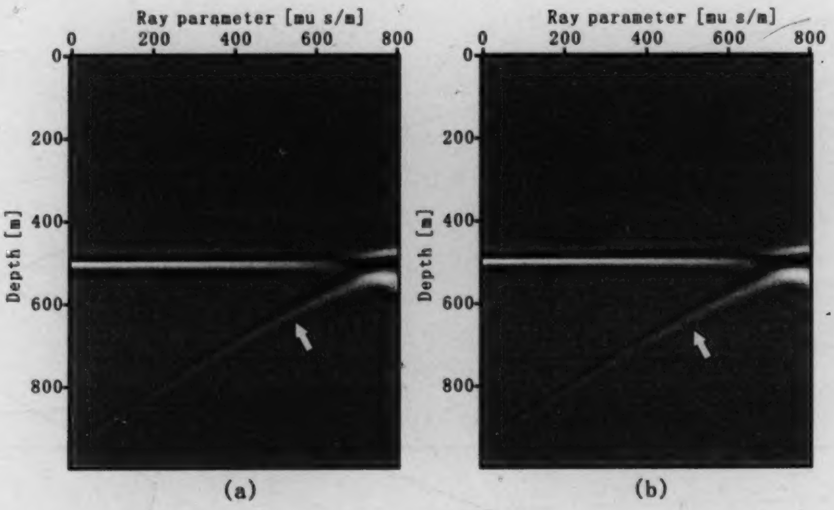
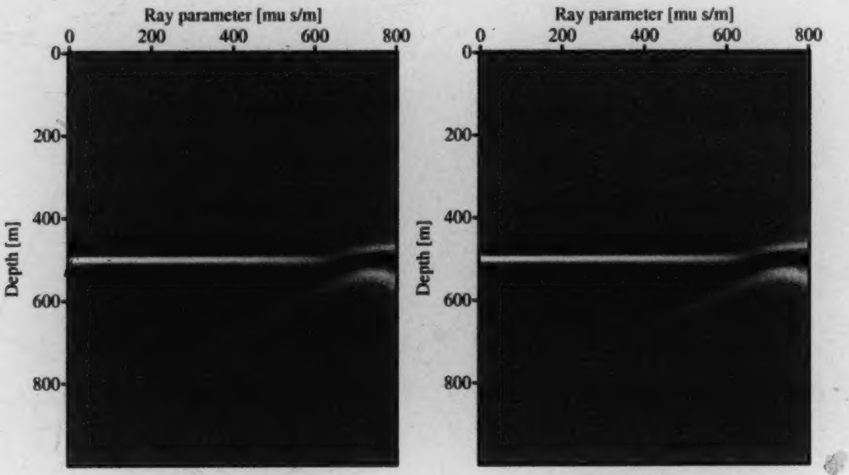


Figure 2.12: Ray parameter dependent image gathers by wave-equation migration. (a) Shot-profile migration. (b) DSR migration. The arrows point the tailing artifacts caused by limited aperture of discrete Fourier transform. These artifacts can be attenuated by tapering and padding zeros (see Figure 2.13)

The tailing artifacts present in the image gathers is a common problem caused by the limited aperture rooted in the discrete Fourier transform. Obviously, if these artifacts are not properly suppressed, the amplitude of deeper events will be disturbed, which hampers further amplitude analysis. One commonly used treatment is to pad zero traces after the large offsets to improve the accuracy of discrete Fourier transform. Theoretically, if infinite zeros are padded, the wrap around effect will disappear (Claerbout, 1985). However, this is too expensive and impractical in application. Another treatment is tapering. In this method, far offsets are muffled gradually, which avoids the sharper change of amplitude along the offset direction. In this thesis, I combine zero padding and tapering to attenuate the high-angle artifacts. For this example, I padded 67 zero traces after the far offsets and apply 20 percent tapering to alleviate the wrap-around effect. The improved image gathers are shown in Figure 2.13



(a) (b)

Figure 2.13: Ray parameter dependent image gathers by wave-equation migration. (a) Shot-profile migration. (b) DSR migration. The tailing artifacts are suppressed by zero padding and tapering.

For rock property studies, the dynamic information, represented by amplitude variation, is very important to the application of seismic data in oil & gas exploration. An imaging algorithm can be especially valuable if it can preserve amplitudes, which are in proportion to the true reflectivities of the subvolume. For this reason, I evaluate the amplitude accuracy of previously described methods by comparing angle dependent amplitude of the relative results. In this toy example, the peak amplitude of the event at depth $z = 500$ m is extracted, normalized and converted to the angle domain using equation 2.120. The curves of angle dependent amplitude by different methods are displayed in Figure 2.14. It can be seen that both shot-profile migration and DSR migration provide accurate amplitude information about interface. Their results are superior to that of the conventional zero-offset imaging method. The performance of zero-offset imaging method can be improved by introducing the amplitude normalization using the source wavefield. The amplitude of the implemented zero-offset image is closer to the theoretical value at medium and high angles. However, the near-offset accuracy deteriorates as a price.

2.8 Common-azimuth approximation

The application of DSR wave-equation operators can be very expensive for 3-D data. The major computational cost lies on multi-dimensional Fourier transforms that convert data back and forth between the space-time domain and the wavenumber-frequency domain. As an example, I analyze the expense of 3-D prestack migration by adopting the zero-offset imaging condition. The data is sampled in five dimensions: inline midpoint (m_x), crossline midpoint (m_y), inline offset (h_x), crossline offset (h_y) and travel time (t). Their dimensions are denoted as N_{m_x}, N_{m_y}, N_{h_x}, N_{h_y} and N_t. First at the surface, the data are converted to the Fourier domain, which involves 5-D forward Fourier transform. Then the wavefield is propagated to each depth using the described DSR downward continuation method. At each depth, before applying the zero-offset imaging condition, the data are converted to the space and temporal frequency domain by 4-D inverse Fourier transform, i.e. change $\tilde{P}(k_{m_x})$, which is repeated for all frequencies. After the imaging condition is applied, the images of all frequencies are summed up to form the final image. To get a sense of the computational cost, I specify the dimension of a 3-D survey that

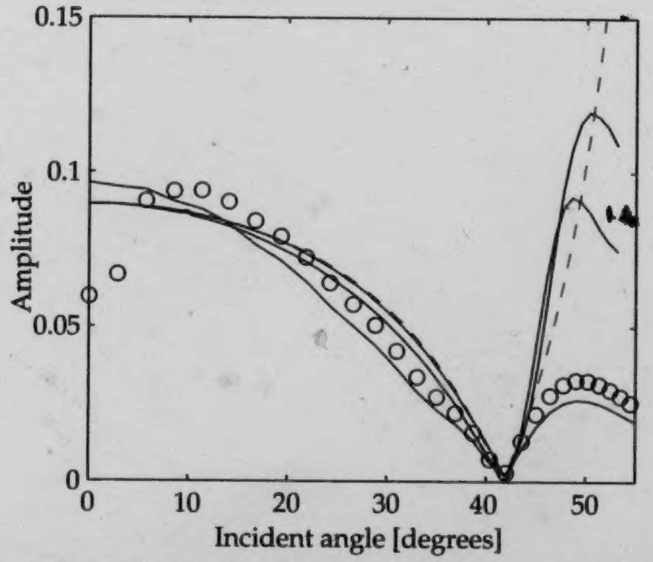


Figure 2.14: AVA curves by different imaging methods. Red dashed: theoretical AVA curve. Blue solid: DSR migration. Black solid: shot-profile migration using plane-wave imaging condition. Green solid: shot-profile migration using zero-offset imaging condition. Black circled: shot-profile migration using normalized zero-offset imaging condition. All amplitudes have been corrected by the inverse of the imaging Jacobian (see Chapter 3).

covers an area of 8 km × 8 km as bellow:

$$\begin{aligned} N_{m_x} &= 400, \\ N_{m_y} &= 400, \\ N_{h_x} &= 24, \\ N_{h_y} &= 16, \\ N_t &= 600, \\ N_\omega &= 300, \\ N_z &= 300, \end{aligned}$$

where N_ω is the Nyquist temporal frequency number, and N_z is the depth number.

At the surface, the 5-D Fourier transform (FT) can be regarded as a chain of 1-D FT in the time domain and 4-D FT in the space domain. The computational time for the transforms using a MIPS R160000 CPU (700 MHz) with the FFTW software (version 2.1.5) is listed as follows:

$$\begin{aligned} 600 \times 1 &: 208.66\mu s \\ 400 \times 400 \times 24 \times 16 &: 92s \end{aligned}$$

Therefore the cost for the surface data conversion is:

$$400 \times 400 \times 24 \times 16 \times 208.66\mu s + 92 \times 300s = 40420s \approx 11 \quad hrs$$

The cost for imaging at all depths is

$$92 \times 300 \times 300 \approx 2280 \quad hrs \tag{2.150}$$

The total cost is about 2291 hours or three months! A real 3-D survey can have tens of such areas, and the cost is prohibitive, which hampers the application of full 3-D DSR migration. To overcome this difficulty, we can resort to two approaches, parallel computing and economical algorithms. The former method entails decomposing a big task into many small tasks that can be done separately in different computers. For example, we can allocate each frequency to a single CPU and calculate the final image by retrieving and summing up the results of all CPUs. The related research is called high performance computing, which is beyond the major interest of this thesis. In this thesis, I use an economical operator, common-azimuth migration (Biondi and Palacharla, 1996), which cuts the computational cost by reducing one dimension of the data, the crossline offset. Only a subset of data with zero crossline offset is utilized to calculate the subvolume wavefields by implementing the DSR operator with a stationary-phase approximation, which greatly cuts the cost. Using this algorithm, about 1/16 of the estimated time is required for the above example.

For notational convenience, common-azimuth data are denoted in the Fourier domain as $\tilde{P}(\omega, k_{m_x}, k_{m_y}, k_{h_x})$. A modelling/migration scheme can be developed by using a recursive formula that links two neighboring depths:

$$\tilde{P}(\omega, k_{m_x}, k_{m_y}, k_{h_x}, z_i) = \tilde{P}(\omega, k_{m_x}, k_{m_y}, k_{h_x}, z_{i-1}) \cdot e^{-ik_z dz}, \tag{2.151}$$

where

$$\begin{aligned} k_z &= DSR(\omega, k_{m_x}, k_{m_y}, k_{h_x}, k_{h_y}, z) \\ &= \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}\left[(k_{m_x} + k_{h_x})^2 + (k_{m_y} + k_{h_y})^2\right]} \\ &+ \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}\left[(k_{m_x} - k_{h_x})^2 + (k_{m_y} - k_{h_y})^2\right]}\,. \end{aligned} \tag{2.152}$$

Since the crossline offset wavenumber k_{h_y} is present in the formula of the vertical wavenumber, it should be eliminated by an integral over the wavenumber for the common-azimuth data. According to the discrete Fourier theory, this procedure means evaluating the wavefield at zero crossline offset. Consequently, the recursive wave propagator is reformulated into

$$\begin{aligned} \tilde{P}(\omega, k_{m_x}, k_{m_y}, k_{h_x}, z_i) &= \int \tilde{P}(\omega, k_{m_x}, k_{m_y}, k_{h_x}, z_{i-1}) e^{-ik_z dz} \qquad (2.153) \\ &= \tilde{P}(\omega, k_{m_x}, k_{m_y}, k_{h_x}, z_{i-1}) \int e^{-ik_z dz} dk_{h_y}, \end{aligned}$$

where the second identity is derived by using the fact that common-azimuth data is independent of the crossline offset wavenumber.The integral is in a typical Fourier form, which can be solved by the stationary-phase approximation (Scales, 1997):

$$\int_{-\infty}^{+\infty} e^{-k_z dz} dk_{h_y} = \sqrt{\frac{2\pi}{-\hat{k}_z'' dz}} e^{-i\hat{k}_z dz + i\frac{\pi}{4}}, \tag{2.154}$$

where $\hat{k}_z$ is the vertical wavenumber k_z evaluated at a stationary point $\hat{k}_{h_y}$ such that

$$\left.\frac{\partial k_z}{\partial k_{h_y}}\right|_{k_{h_y}=\hat{k}_{h_y}} = 0, \tag{2.155}$$

$$\left.\frac{\partial^2 k_z}{\partial k_{h_y}^2}\right|_{k_{h_y}=\hat{k}_{h_y}} \neq 0 \tag{2.156}$$

Solving equation 2.155 leads to

$$\hat{k}_{h_y}(z) = k_{m_y}\frac{\sqrt{\frac{1}{v^2}-\frac{1}{4\omega^2}(k_{m_x}+k_{h_x})^2}-\sqrt{\frac{1}{v^2}-\frac{1}{4\omega^2}(k_{m_x}-k_{h_x})^2}}{\sqrt{\frac{1}{v_m^2}-\frac{1}{4\omega^2}(k_{m_x}+k_{h_x})^2}+\sqrt{\frac{1}{v^2}-\frac{1}{4\omega^2}(k_{m_x}-k_{h_x})^2}} \tag{2.157}$$

For stability, the amplitude term in equation 2.154 is ignored. This is reasonable since the cumulative product of a number larger than one will blow up the wavefield. On the other hand, a number smaller than one will decimate the wavefield with the propagation of the wavefield. Actually the amplitude term is used to convert the amplitude of recorded common-azimuth data with point sources to 2-D amplitude with line sources (Biondi, 2003a). To verify this statement, I consider the situation of $V(z)$[1] media. In this case, the crossline offset wavenumber disappears. The vertical wavenumber k_z degrades into a 2-D version DSR operator. Therefore, for amplitude fidelity the amplitude of seismic data should be corrected to before applying common-azimuth wave-equation migration. For constant-velocity media, the geometrical spreading factor is the inverse of the wave-propagation distance for point sources and the square root of the distance for line sources (Berkhout, 1984). It is clear that multiplication with the square root of the distance or travel time approximately converts 3-D amplitudes to 2-D amplitudes. This is often used in the industry even when the constant-velocity assumption is violated.

By ignoring the amplitude correction term, the downward continuation formula for 3-D common-azimuth data is reformulated as follows:

$$\tilde{P}(\omega, k_{m_x}, k_{m_y}, k_{h_x}, z_i) = \tilde{P}(\omega, k_{m_x}, k_{m_y}, k_{h_x}, z_{i-1}) \cdot e^{-ik_z dz}, \tag{2.158}$$

where

$$k_z = DSR(\omega, k_{m_x}, k_{m_y}, k_{h_x}, \hat{k}_{h_y}, z), \tag{2.159}$$

---

[1] Velocity only changes in the vertical direction

where $\hat{k}_{h_y}$ is given by equation 2.157.

The common-azimuth approximation provides an efficient tool for seismic migration and modelling. In migration, the surface wavefield is backward propagated to each depth by downward continuation. Then at each depth AVP common image gathers can be calculated by applying the radial-trace transform. Since only inline offset wavenumber is available, the resulting imaging gathers are in the inline offset ray parameter domain. The image gathers can be converted to apparent angle domain in the inline direction using equation 2.120. If the crossline structural dip can be ignored, the angle is readily the true incident angle. On the other hand, when the crossline wavenumber k_{m_y} is not zero, the dispersion relationship defined by the vertical wavenumber deviates from the 2-D case. The larger is the crossline structural dip, the larger is the error. The common-azimuth modelling starts from each depth where the adjoint of radial-trace transform is applied to the AVP common image gathers to calculate local wavefield. Then the wavefield is upward continued to the surface. At the surface, waves arising from all depths are summed up to simulate the seismic response.

Although common-azimuth migration is not the exact solution of full 3-D wave equation, it is better than 2-D migration since it uses crossline information. It can be regarded as the chaining of inline 2-D prestack migration and crossline poststack migration (Biondi, 2003a).

The strategy of common-azimuth approximation is especially useful for marine data, which usually have a narrow-azimuth geometry. The data can be converted to a common azimuth without losing much accuracy (Biondi et al., 1998). Furthermore, the method may also be beneficial for land data. One possible application is to apply the imaging algorithm to data of various azimuths and study the amplitude variation with angle and azimuth. The analysis of Amplitude Versus Angle and aZimuth (AVAZ) can be used to measure fracture density and fracture strike within carbonates (Rüger, 1996; Alhavas et al., 2003; Gray and Head, 2000).

2.9 Implementation for laterally variant media

By far, wave-equation style operators have been introduced for laterally invariant media. In the real world, the laterally invariance assumption is almost always violated. Therefore the solution is usually unsatisfactory due to the operator mismatch. Many geophysicists have attempted to improve the operator accuracy by numerical methods.

One straight-forward wavefield calibration is to employ interpolation technique to fit lateral velocity variations (Gazdag, 1984). The method is called *Phase Shift Plus Interpolation* or abbreviated as PSPI. The method entails downward continuing the wavefield with some reference velocities in the wavenumber-frequency domain, inverse Fourier transform to the space-frequency domain and interpolation in the space-frequency domain. With proper choice of reference velocities, the algorithm can effectively image complex media with strong velocity variations. The forward Fourier transform and inverse Fourier transform are applied to the wavefields in respect to different reference velocities. As a result, the computational cost is multiplied compared with the simple phase-shift migration for $V(z)$ media (Gazdag, 1978).

For efficiency, when the velocity variation is slight, Split-step correction (Stoffa et al., 1990; Popovici, 1996) is preferred. In this method the wave continuation is conducted in two steps. In the first step, the wavefield is phase-shifted using a constant macro velocity in the wavenumber-frequency domain. In the second step, velocity correction is applied in the space-frequency domain. Similar to PSPI, The accuracy of Split-step methods can be enhanced by an interpolation approach called *Extended Split-step Fourier Migration* (ESFM) (Kessinger, 1992) or split-step PSPI. This method is also robust in handling large lateral velocity variations.

As shown further bellow, the split-step method attempts to correct the first order error that is brought by the velocity mismatch. An alternative way to further improve split-step migration is to decrease higher order errors. For example, Fourier Finite Difference (FFD) method (Ristow and Rühl, 1994) splits the problem into three steps. The first step is the same as the well-known phase-shift method, propagating the wavefield in the wavenumber-frequency domain using a constant background velocity. The second step is to apply split-step correction in the space-frequency domain. The third step is to correct higher order wavenumber errors in the space-frequency domain by finite difference numerical methods. Similar strategy is adopted in pseudo-screen methods (Xie and Wu, 1998; Huang and Wu, 1996; Jin et al., 2002). One drawback of FFD and pseudo-screen methods is the frequency dispersion artifacts brought by the finite difference step. Biondi (2002) proposed an efficient interpolation method to address this issue.

Since finite difference methods are not easy to implement for iterative inversion, I adopt split-step and extended split-step corrections for different levels of lateral velocity variations. The detail of these two methods is described as bellow. For completeness, I also show the scheme of PSPI migration. Some tests of these operators on the 2-D poststack SEG/EAGE salt data are shown in Chapter 4. According to the tests, the image quality given by PSPI and extended split-step methods are very similar.

Split-step Fourier migration

The accuracy of DSR wave equation migration is dependent on the calculation of the two single square roots (see equation 2.137), or vertical wavenumbers in the Fourier domain. One big challenge is the lateral velocity variation. The general problem of velocity correction can be posed as bellow in the temporal frequency and space domain:

$$SSR(\omega, v(\mathbf{m}), \mathbf{m}) = SSR(\omega, c(\mathbf{m}), \mathbf{m}) + \Delta SSR, \tag{2.160}$$

where SSR denotes the single square root, $v(\mathbf{m})$ is the true velocity at midpoint $\mathbf{m}$, and $c(\mathbf{m})$ is the reference velocity or the macro velocity, which is usually the average velocity at the current depth where wave is propagated. For simplicity, I refer to the formula for the 2-D case given by Ristow and Rühl (1994) to account for lateral velocity variations:

$$\sqrt{\frac{\omega^2}{v^2} + \frac{\partial^2}{\partial x^2}} = \underbrace{\sqrt{\frac{\omega^2}{c^2} + \frac{\partial^2}{\partial x^2}}}_{Term1} + \underbrace{(\frac{\omega}{v} - \frac{\omega}{c})}_{Term2} + \underbrace{\frac{\omega}{v}(1 - \frac{c}{v})\frac{\frac{v^2}{\omega^2} \cdot \frac{\partial^2}{\partial x^2}}{a + b \cdot \frac{v^2}{\omega^2} \cdot \frac{\partial^2}{\partial x^2}}}_{Term3}, \tag{2.161}$$

where

$$\begin{aligned} a &= 2, \\ b &= \frac{1}{2}\Big[\frac{c^2}{v^2} + \frac{c}{v} + 1\Big]. \end{aligned}$$

The expression is derived by analyzing the Taylor series of the single square root. The calculation of the single root can be split into three steps in honor of the three terms respectively. The first step is usually done in the wavenumer-frequency domain. Applying the spatial Fourier transform leads to

$$k_{zc} = \sqrt{\frac{\omega^2}{c^2} - k_x^2}. \tag{2.162}$$

It is clear that this is the vertical wavenumber for phase shift migration using the constant reference velocity. The second term is the first order correction, which is done in the space and temporal frequency domain. The third term is the higher order correction, which is done by finite difference methods in the space and temporal frequency domain.

The standard Split-step Fourier migration only takes into account the first two terms of equation 2.161. The algorithm splits the velocity corrections into two steps as discussed earlier. The processing flow of the standard split-step Fourier migration is portrayed in Figure 2.15.The scheme can be further implemented for the DSR operator, which involves correcting velocity variations in two single squares roots (Popovici, 1996). These two square roots are related to the source and the receiver respectively. Therefore the split-step correction is conducted at two locations, $\mathbf{x} + \mathbf{h}$ for the receiver and $\mathbf{x} - \mathbf{h}$ for the source.

The extended split-step Fourier migration (Kessinger, 1992) is based on the standard split-step migration. The method applies split-step migrations using several reference velocities instead of only one constant velocity. Several copies of propagated wavefields calculated by these reference velocities are interpolated to match local velocities. The interpolation is conducted in the space and temporal frequency domain. Figure 2.16 displays the scheme of the extended split-step Fourier migration using two reference velocities c_1 and c_2. Since separate interpolation of phase and amplitude may lead to instability (Etgen, 1994), linear interpolation of complex numbers is adopted. For example, if we know that the velocity wavefield of a location is $v(x, z)$, then the wavefield at this point is calculated by a linear interpolation of those propagated wavefields using two neighboring reference velocities c_1 and c_2 ($c_1 < c_2$):

$$\begin{aligned} \alpha &= \frac{c_2 - v}{c_2 - c_1}, \\ \tilde{P}(x, v, z) &= \alpha \cdot \tilde{P}(x, c_1, z) + (1 - \alpha) \cdot \tilde{P}(x, c_2, z), \end{aligned}$$

where $\tilde{P}$ denotes the wavefield in the wavenumbe and temporal frequency domain.

Phase Shift Plus Interpolation migration

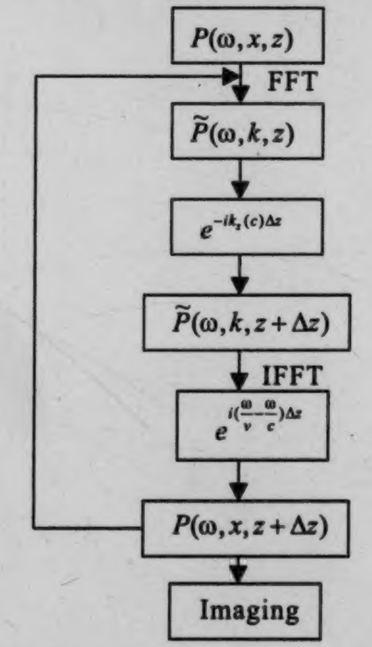

$P(\omega,x,z)$
FFT
$\widetilde{P}(\omega,k,z)$
$e^{-ik_z(c)\Delta z}$
$\widetilde{P}(\omega,k,z+\Delta z)$
IFFT
$e^{i(\frac{\omega}{v}-\frac{\omega}{c})\Delta z}$
$P(\omega,x,z+\Delta z)$
Imaging

Figure 2.15: The scheme of the split-step Fourier migration. FFT: Fast Fourier Transform. IFFT: Inverse Fast Fourier Transform.

PSPI migration (Gazdag, 1984) is similar to the extended split-step Fourier migration by using multiple reference velocities to interpolate the wavefields. For each reference velocity c, the split-step correction is rearranged into

$$\sqrt{\frac{\omega^2}{v^2}+\frac{\partial^2}{\partial x^2}} \approx \underbrace{\frac{\omega}{v}}_{Term1} + \underbrace{(\sqrt{\frac{\omega^2}{c^2}+\frac{\partial^2}{\partial x^2}}-\frac{\omega}{c})}_{Term2} \tag{2.163}$$

The first term is applied in the space and temporal frequency domain, and the second term is applied in the wavenumber and temporal frequency domain. Explicitly, the second step can be expressed as

$$k_{zc} - k_0 = \sqrt{\frac{\omega^2}{c^2} - k_x^2} - \frac{\omega}{c}. \tag{2.164}$$

Figure 2.17 explains the scheme of PSPI using two reference velocities c_1 and c_2. Since

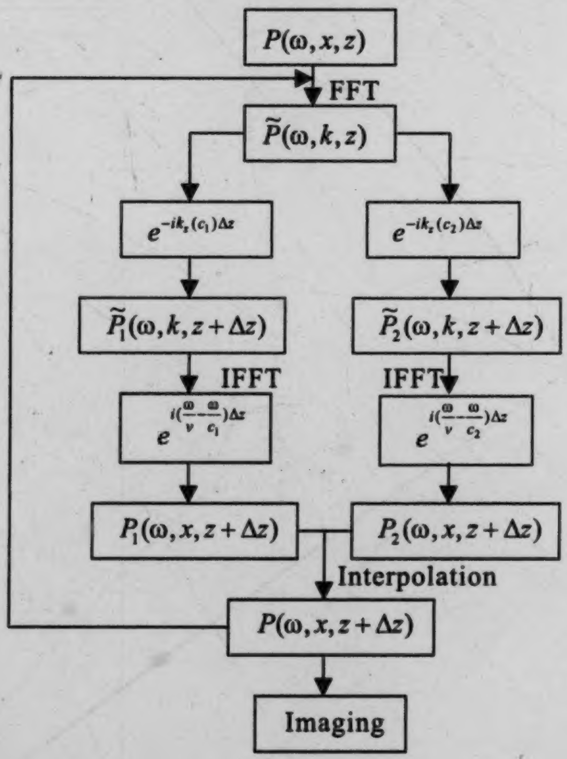

$P(\omega,x,z)$
FFT
$\tilde{P}(\omega,k,z)$
$e^{-ik_z(c_1)\Delta z}$
$e^{-ik_z(c_2)\Delta z}$
$\tilde{P}_1(\omega,k,z+\Delta z)$
$\tilde{P}_2(\omega,k,z+\Delta z)$
IFFT
IFFT
$e^{i(\frac{\omega}{v}-\frac{\omega}{c_1})\Delta z}$
$e^{i(\frac{\omega}{v}-\frac{\omega}{c_2})\Delta z}$
$P_1(\omega,x,z+\Delta z)$
$P_2(\omega,x,z+\Delta z)$
Interpolation
$P(\omega,x,z+\Delta z)$
Imaging

Figure 2.16: The scheme of the extended split-step Fourier migration (split-step PSPI). FFT: Fast Fourier Transform. IFFT: Inverse Fast Fourier Transform.

separate interpolation of phase and amplitude may lead to instability (Etgen, 1994), linear interpolation of complex wavefields is used to fit the local wave speed.

Optimization of reference velocities

In the extended split-step (split-step PSPI) and PSPI migration, proper reference velocities should be chosen to provide satisfactory results. It is obvious that for complex media, more reference velocities lead to higher accuracy, but also higher computational cost. To balance between accuracy and efficiency, we should optimize the reference velocities to adequately represent all velocities. One idea is to analyze the distribution of the reference velocities so that we can estimate some representative velocities, which most velocities cluster around. This can be done numerically by using the entropy as the

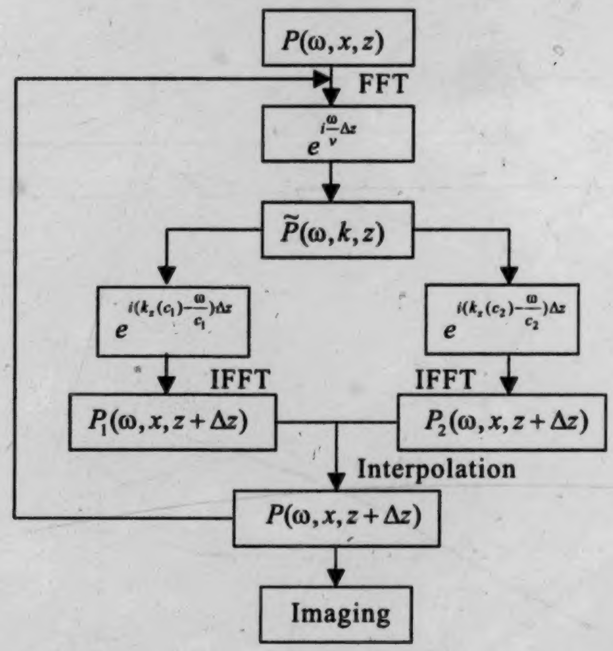


Figure 2.17: The scheme of the PSPI migration. FFT: Fast Fourier Transform. IFFT: Inverse Fast Fourier Transform.

measurement of the velocity dispersion (Bagaini et al., 1995).

The entropy is defined as

$$s = -\sum_{k=1}^{L} p_k \ln p_k, \tag{2.165}$$

where p_k is the probability of the k th event. The entropy can be used to evaluate the simplicity of the distribution of a variable. The smaller is the entropy, the simpler is the distribution. Taking the migration velocity as an example, we like the velocity structure to be simple since only a few reference velocities are necessary for PSPI methods. The easiest case is constant velocity. Then the entropy is zero. For non-constant velocity model, we can study the velocity distribution by counting the number of velocities that locate in the prepared velocity bins and then calculating the probability that velocity locates in each specific velocity bin. For simplicity, at each depth of wave extrapolation, the velocity bins are prepared with a constant spacing, $\Delta c = (c_{max} - c_{min})/L$, where L is the total bin number, c_{max} is the maximum velocity, and c_{min} is the minimum velocity. Therefore, the velocity bins are: $[c_1 \quad c_2], [c_2 \quad c_3], ..., [c_L \quad c_{L+1}]$, where $c_{k+1} - c_k = \Delta c$. The probability for these bins are denoted as $p_k, \quad k = 1, L$.

The optimal velocity number can be estimated by the following formula:

$$N_c = int(e^s), \tag{2.166}$$

where *int* means take the closest integer number. It is easy to verify that the velocity number is one for constant velocity model and L for the worst case when all bins have the same probability $1/L$. After calculating the reference velocity number N_c, we can estimate the optimal reference velocities such that

$$\frac{i}{N_c} = \int_{c_{min}}^{v_{ref}^{i+1}} p(v)dv, \qquad i = 1, ..., N_c, \tag{2.167}$$

where v_{ref}^{i+1} is the $(i + 1)$ th reference velocity. It is clear that the anticipated optimal reference velocities satisfy the following relationships:

$$\begin{aligned} p([v_{ref}^1 \quad v_{ref}^2]) &= \frac{1}{N_c}, \\ p([v_{ref}^2 \quad v_{ref}^3]) &= \frac{1}{N_c}, \\ &\dots \\ p([v_{ref}^{N_c} v_{ref}^{N_c+1}]) &= \frac{1}{N_c}, \end{aligned} \tag{2.168}$$

with an extra reference velocity $v_{ref}^{N_c+1}$ defined.

The reference velocities are calculated in an iterative way as follows

$$\begin{aligned} v_{ref}^1 &= c_{min}, \\ v_{ref}^{i+1} &= c_k + \left[\frac{i}{N_c} - y_k\right] \frac{c_{k+1} - c_k}{y_{k+1} - y_k}, \qquad i = 1, ..., N_c, \end{aligned} \tag{2.169}$$

where k is the index such that $y_k < \frac{i}{N_c} < y_{k+1}$, where y_k is a probability defined as

$$y_k = \int_{c_{min}}^{c_k} p(v)dv, \qquad k = 2, ..., L. \tag{2.170}$$

Equation 2.169 can be understood as an interpolation in the original velocity bin $[c_k \quad c_{k+1}]$ using the difference of cumulative velocity probabilities. Figure 2.18 shows the distribution of velocities of the 2-D SEG/EAGE model at depth $z = 6520ft$. It is clear that the calculated optimal velocities represent most of the velocities. However, there are still some missing velocities around $9000ft/m$. Better strategy of velocity optimization should be able to detect such velocity of local probability extrema.

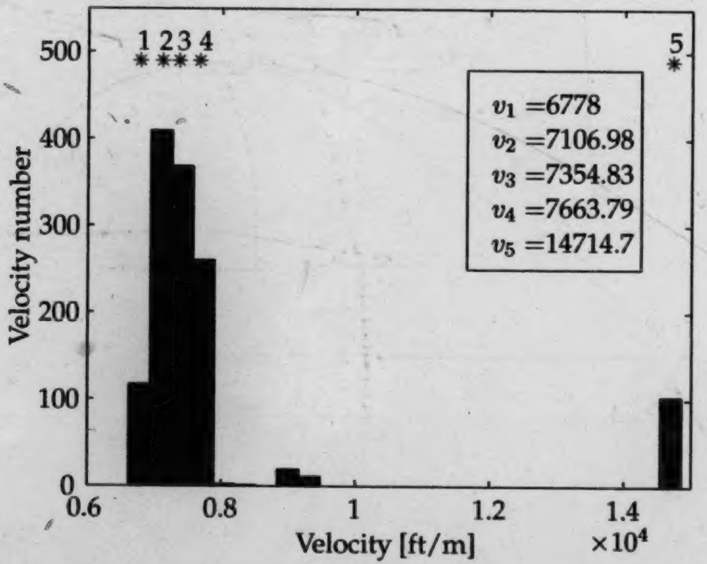


Figure 2.18: Velocity distribution of the 2-D SEG/EAGE salt model at depth $z = 2520ft$. The red stars represent the calculated optimal reference velocities for PSPI migration.

2.10 Summary

In this chapter, I have gone through theoretical aspects of seismic modelling and migration including ray based and wave equation based methods. Especially, an economical strategy for f 3-D AVP modelling and migration is derived by adopting the common azimuth approximation. Almost all the machinery is now ready for the iterative inversion that will be discussed in the next chapter. To close this chapter, I will generalize the forward modelling and the adjoint operators in matrix form. This will bring a Linear Algebra *flavour* to our problem, and facilitate the description of the numerical inverse problem tackled in Chapter 3. The modelling operator **L** is defined as:

$$\mathbf{L}: \qquad \mathbf{d} = \mathbf{W}_s \mathbf{L}_\omega \mathbf{A} \mathbf{m}, \tag{2.171}$$

where **d** is the seismic data in 3-D common-azimuth volume, $\mathbf{W}_s$ is the source signature that is added in the frequency domain, $\mathbf{L}_\omega$ is the wave propagator, which can be the split-step or split-step PSPI wave extrapolator accounting for lateral velocity variations, **A** is the adjoint of radial trace transform, and **m** is the 3-D common-azimuth AVP image gathers. The operators are applied on the fly from the right to the left for each single frequency. The resulting wavefields are accumulated to produce the final seismic data.

The adjoint operator $\mathbf{L}'$, or AVP migration, can be cast as:

$$\mathbf{L}': \qquad \tilde{\mathbf{m}} = \mathbf{A}' \mathbf{L}'_\omega \mathbf{W}'_s \mathbf{d}, \tag{2.172}$$

where the tildes symbolize the adjoints of the operators used in the seismic modelling. Again, the operators are applied in a row from the right to the left, which is opposite to the order of seismic modelling. Since the source signature, especially its phase, is difficult to estimate, it is usually not considered in the iterative inversion, which implicitly limits the resolution due to the wavelet mismatch. This is discussed later in Chapter 4.

Chapter 3

Regularized common-azimuth least-squares AVP migration

3.1 Introduction

Posing migration as an regularized inverse problem can lead to high resolution solutions (Ehinger and Lailly, 1991; Nemeth et al., 1999; Kuehl and Sacchi, 1999; Duquet et al., 2000; Prucha and Biondi, 2002). These methods can provide promising results since they can fit the seismic data and at the same time impose specific properties on the solution. This features are consistent with *a priori* information about the model. This strategy has two benefits. First, by fitting the seismic data, we try to invert the linearized earth system defined by some approximate transforms. These transforms are usually band-limited and they provide limited information of the earth model. During inversion, this drawback can be alleviated by recovering higher-frequency attributes. Furthermore the geometry of data acquisition can be encoded in the operator. This is fundamental since the data acquisition footprint can be removed by incorporating a data covariance matrix (Tarantola, 1987). Second, our knowledge of the model is taken into account by regularizing the solution, which essentially narrows down the family of feasible models.

In the geophysical community, the idea of constrained or regularized inversion has been widely used in seismic data processing. These methods vary according to the problem at hand and the adopted approaches to regularization. For example, the classical deconvolution problem inverts the convolution matrix to reconstruct the reflectivity model from band-limited seismograms. In the linear deconvolution problem, the minimum norm assumption (regularization) of the reflectivity model is often applied to stabilize the solution and suppress the noise. As shown in the text bellow, in the explicit form of the least-squares solution, this is accomplished by introducing pre-whitening or damping in the solution. Although the linear inversion improve the solution, high frequency information in the original model can not be retrieved. On the other hand, non-quadratic regularization methods can be used to attained high resolution solutions. For example, by assuming sparseness (sparse reflectivity assumption), one can use sparse regularization to improve the resolution of solution. Among these sparse regularization methods we can cite: regularization with the l_1 norm (Claerbout, 1985; Taylor et al., 1979), and Cauchy norm (Sacchi, 1997) and, the well-known minimum entropy deconvolution where the solution is obtained by maximizing the varimax norm (Wiggins, 1978).

Obviously, migration is more difficult than deconvolution. First of all, migration is more expensive due to the model and data dimensions. Especially, the DSR wave-equation migration inverts the full volume of data simultaneously, which requires large memory and high computational cost. Second, the problem does not have an explicit matrix form for the operator, the wave equation based migration usually cannot be expressed in matrix form or even in a closed mathematical formula. However, it can be decomposed into a cascade of linear operators.

In this chapter, I first describe regularized least-squares migration. Then an iterative inversion method based in the method of Conjugate Gradients (CG) is described. Both quadratic and non-quadratic regularization methods are discussed. Simple test examples are provided to examine the validness of these methods. Their application in migration is tested in Chapters 4 and 5.

3.2 Regularized least-squares migration

Consider seismic data as the result of a linear transformation on an earth model **m**

$$\mathbf{d} = \mathbf{L}\mathbf{m} + \mathbf{n}, \tag{3.1}$$

where **d** denotes the pre-processed seismic data, **L** is the common-azimuth forward operator described in the previous chapter, **m** is the earth model, a set of ray-parameter dependent common image gathers (CIGs), and $\mathbf{n}$ denotes additive noise. Conventional migration entails applying $\mathbf{L}'$, the adjoint of $\mathbf{L}$, to the observed data. The adjoint operator $\mathbf{L}'$ can be decomposed into two steps: wavefield extrapolation and ray-parameter imaging (full detail is given in Chapter 2). The second step amounts to a change of variables by a radial-trace transform (Sava et al., 2001). When the data are properly sampled, the amplitude in the CIG can be corrected by applying an approximate inverse of the imaging Jacobian. This implemented method is called true-amplitude migration. The theory is briefly introduced at the end of this chapter. The Jacobian weighting attempts to make the adjoint operator behave like the inverse. However, this correction is not sufficient to achieve good amplitude fidelity in situations where the image is corrupted by aliasing artifacts introduced by inadequate spatial sampling. These artifacts can be alleviated, however, by minimizing a cost function of the form:

$$F(\mathbf{m}) = ||\mathbf{W}(\mathbf{d} - \mathbf{L}\mathbf{m})||^2 + \lambda^2 R(\mathbf{m}), \tag{3.2}$$

where $\mathbf{W}$ is a diagonal weighting matrix (data-space weighting) used to decrease the influence of missing observations in the migrated image. The diagonal elements of $\mathbf{W}$ consist of zeros and ones, weighting dead traces and live traces, respectively. R is the regularization function, and λ is a trade-off parameter that controls the amount of regularization.

The regularization operator R can be defined in different ways:

- $R(\mathbf{m}) = ||\mathbf{m}||_2^2$, the quadratic norm of the model. The solution is called the minimum quadratic norm solution.
- $R(\mathbf{m}) = ||\mathbf{D}_{1h_x}\mathbf{m}||_2^2$, also a quadratic norm, where $\mathbf{D}_{1h_x}$ is the first order derivative operator applied along the offset ray parameter direction. Since the derivative is a high-pass filter, minimizing the cost function is identical to penalizing high-frequency solutions. As a result, the final solution is a smooth solution.
- $R(\mathbf{m}) = \sum_{i=1}^{M} |m_i|$, the l_1 norm. M is the length of the model vector. This regularization introduces sparseness in the solution.
- $R(\mathbf{m}) = \sum_{i=1}^{M} \ln(1 + m_i^2/\sigma_m^2)$, the Cauchy norm. M is the length of the model vector. The implemented cost function also provides a sparse solution.

The above cost functions can be generalized under a Bayesian scheme by assuming different *a priori* distributions of the unknown model parameters (Youzwishen, 2001). The two aforementioned quadratic regularization terms can be derived by assuming a Gaussian *a priori* distribution of the model or the derivative of the model. The cost function for non-quadratic regularization for sparse solutions can be derived by assuming an exponential or Cauchy *a priori* distribution of the model parameters.

The optimization problem entails solving:

$$\frac{dF(\mathbf{m})}{d\mathbf{m}} = 0. \tag{3.3}$$

The respective solutions of the above cost functions are listed as follows:

- $\mathbf{m} = (\mathbf{L}'\mathbf{W}'\mathbf{W}\mathbf{L} + \lambda^2\mathbf{I})^{-1}\mathbf{L}'\mathbf{d}_{obs}$, where $\mathbf{d}_{obs}$ is the observed incomplete data and $\mathbf{I}$ is an identity matrix. This is the classic Damped Least-Squares (DLS) solution. The trade-off parameter can also be called pre-whitening coefficient which is used to stabilize the solution and to attenuate noise.

- $\mathbf{m} = (\mathbf{L}'\mathbf{W}'\mathbf{W}\mathbf{L} + \lambda^2\mathbf{D}'_{1h_x}\mathbf{D}_{1h_x})^{-1}\mathbf{L}'\mathbf{d}_{obs}$. This is a smooth solution. The trade-off parameter λ controls the amount of smoothness.

- $\mathbf{m} = (\mathbf{L}'\mathbf{W}'\mathbf{W}\mathbf{L} + \lambda^2\mathbf{Q}(\mathbf{m}))^{-1}\mathbf{L}'\mathbf{d}_{obs}$, where $\mathbf{Q}$ is a diagonal matrix defined by

  $$Q_{ii} = \begin{cases} |m_i|^{-1}, & if|m_i| > \epsilon \\ \epsilon^{-1}, & if|m_i| < \epsilon \end{cases}, \tag{3.4}$$

  where Q_{ii} is the i th diagonal element, and ϵ is a threshold value to detect sparseness.

- $\mathbf{m} = (\mathbf{L}'\mathbf{W}'\mathbf{W}\mathbf{L} + \lambda^2\mathbf{Q}(\mathbf{m}))^{-1}\mathbf{L}'\mathbf{d}_{obs}$, where $\mathbf{Q}$ is a diagonal matrix defined by

  $$Q_{ii} = \frac{1}{1 + m_i^2/\sigma_m^2}, \tag{3.5}$$

  where Q_{ii} is the ith diagonal element, and σ_m is a scaling factor.

It can be seen that the first two problems are linear, in other words, the use of a quadratic cost function has led to a linear system of equations. The major task is to invert the two Hessian matrices, which are independent of the model. On the other hand, the solution with sparse regularization is non-linear since the expression contains the solution itself.

The non-linear problem can be solved iteratively by solving a series of linear problems. In each linear problem, the model-dependent diagonal weighting matrix is estimated from the previous solution. Since the diagonal weighting matrix is updated gradually, the algorithm is called Iterative Reweighted Least-Squares (IRLS) method (Scales and Smith, 1994). The solution can be expressed as $\mathbf{m}_i = (\mathbf{L}'\mathbf{W}'\mathbf{W}\mathbf{L} + \lambda^2\mathbf{Q}(\mathbf{m}_{i-1}))^{-1}\mathbf{L}'\mathbf{d}_{obs}$, where $\mathbf{m}_i$ and $\mathbf{m}_{i-1}$ are the solutions at the i th and $(i-1)$ th iteration, respectively.

The least-squares solution can be understood as a de-blurred version of the adjoint solution. To make this clear, I denote the adjoint solution $\mathbf{L}'\mathbf{d}_{obs}$ as $\tilde{\mathbf{m}}$. Then the least-squares solution can be reformulated as

$$\mathbf{m} = \mathbf{H}^{-1}\tilde{\mathbf{m}}, \tag{3.6}$$

where the Hessian matrix $\mathbf{H}$ varies with different regularization. Therefore, the essence of least-squares inversion is to calculate the inverse of the Hessian matrix and use it to filter the adjoint solution.

In this thesis, I choose two kinds of regularization methods. One is the smooth regularization using the quadratic norm of the model derivative, and the other is the sparse regularization using the Cauchy norm. The smoothness is forced along the offset ray parameter direction, and the sparseness is applied in the depth direction of the stacked image. This is reasonable since the amplitude variation versus ray parameter is smooth, and the theoretical full-band reflectivity series in the vertical direction should be sparse.

As explained earlier, the direct inversion of large ill-posed problems is difficult and expensive. Therefore, in this thesis I use an efficient iterative inversion method, Conjugate Gradients (CG) (Hestenes and Steifel, 1952), to solve these problems.

3.3 Conjugate gradients algorithm

The original CG algorithm (Hestenes and Steifel, 1952) attempts to solve the following system of equations:

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{3.7}$$

where $\mathbf{A}$ is a positive-definite symmetric matrix, $\mathbf{x}$ is a model vector, and $\mathbf{b}$ is a data vector. However a typical seismic inversion problem is to solve a rectangular, often underdetermined, system like

$$\mathbf{Lm} = \mathbf{d}. \tag{3.8}$$

The latter can be reformulated for the standard CG algorithm

$$\mathbf{L'Lm} = \mathbf{L'd}, \tag{3.9}$$

by recognizing that $\mathbf{L'L}$ is a symmetric matrix and assuming it is positive definite. A modified CG algorithm called CGLS (Scales, 1987), generalized further below, can solve the problem without constructing the matrix $\mathbf{L'L}$.

The regularized least-squares migration is more complicated due to the data weighting matrix and the regularization operator. To apply the CGLS algorithm, the system can be expressed in augmented matrix form. For example, the smooth regularization problem can be recast as:

$$\begin{pmatrix} \mathbf{WL} \\ \lambda \mathbf{D}_{1h_x} \end{pmatrix} \cdot \mathbf{m} = \begin{pmatrix} \mathbf{d} \\ \mathbf{0} \end{pmatrix} \tag{3.10}$$

Denoting the left side augmented matrix as $\tilde{\mathbf{W}}$ and the right side augmented data vector as $\tilde{\mathbf{d}}$ yields

$$\tilde{\mathbf{L}}\mathbf{m} = \tilde{\mathbf{d}}, \tag{3.11}$$

which is the standard inversion problem.

The CGLS algorithm for the problem 3.8 is generalized as bellow:

First the model is initiated with zeros, and the data vector is initiated with the observation, i.e., $\mathbf{m}_0 = \mathbf{0}$, $\mathbf{s}_0 = \mathbf{d}$. Two supplementary model vectors are calculated by the adjoint operation: $\mathbf{r}_0 = \mathbf{p}_0 = \mathbf{L}'\mathbf{s}_0$. A data vector $\mathbf{q}$ is started with: $\mathbf{q}_0 = \mathbf{L}\mathbf{p}_0$. Then repeat the following steps for $i = 0, 1, 2...$

$$\begin{aligned}
\alpha_{i+1} &= \frac{\mathbf{r}_i \cdot \mathbf{r}_i}{\mathbf{q}_i \cdot \mathbf{q}_i} \\
\mathbf{m}_{i+1} &= \mathbf{m}_i + \alpha_{i+1}\mathbf{p}_i \\
\mathbf{s}_{i+1} &= \mathbf{s}_i - \alpha_{i+1}\mathbf{q}_i \\
\mathbf{r}_{i+1} &= \mathbf{L}'\mathbf{s}_{i+1} \\
\beta_{i+1} &= \frac{\mathbf{r}_{i+1} \cdot \mathbf{r}_{i+1}}{\mathbf{r} \cdot \mathbf{r}} \\
\mathbf{p}_{i+1} &= \mathbf{r} + \beta\mathbf{p}_i \\
\mathbf{q}_{i+1} &= \mathbf{L}\mathbf{p}_{i+1}.
\end{aligned}$$

3.4 Least-squares migration with smooth regularization

For convenience, I denote the least-squares migration with smooth regularization as RLSM and the least-squares migration with sparse regularization as SLSM. As analyzed previously, the problem can be reformulated in the augmented matrix form 3.11. The CG algorithm reduces to the sequential application of the following operators: migration $\mathbf{L}'$, de-migration $\mathbf{L}$, un-smoothing $\mathbf{D}_{1\,h_x}$, and the adjoint of un-smoothing $\mathbf{D}'_{1h_x}$. If we express the data vector in two parts

$$\mathbf{d} = \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix}, \tag{3.12}$$

then the modelling procedure $\tilde{\mathbf{L}}$ can be regarded as two steps:

$$\mathbf{d}_1 = \mathbf{WLm}, \tag{3.13}$$

$$\mathbf{d}_2 = \lambda \mathbf{D}_{1h_x}\mathbf{m}. \tag{3.14}$$

The adjoint operator $\tilde{\mathbf{L}}'$ is the combination of two adjoints:

$$\mathbf{m} = \mathbf{L}'\mathbf{W}'\mathbf{d}_1 + \lambda \mathbf{D}'_{1h_x}\mathbf{d}_2. \tag{3.15}$$

The operator $\mathbf{D}_{1h_x}$ is a discrete derivative (high-pass filter), and the transpose operator $\mathbf{D}'_{1h_x}$ is a discrete negative derivative (Claerbout, 2004). For example, if the length of ray parameter image gather at each depth of the common image gather is four, then the forward derivative operator is

$$\mathbf{D}_{1h_x} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

and the adjoint is

$$\mathbf{D}'_{1h_x} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

The choice of trade-off parameter λ in equation 3.14 and 3.15 can pose a challenge for a large-scale linear problem like regularized least-squares migration. It is obviously not feasible to determine λ from a trade-off curve (L curve method (Hansen, 1998)). In practice, I therefore iterate the following procedure until a good value for λ is found. I run least-squares migration for a couple of iterations for only a few frequencies starting with a small trade-off parameter, for example $\lambda = 0.0001$, and monitor the data misfit of a few common midpoint (CMP) gathers. If, within the first two iterations, the data residual norm has been reduced to 30-60 percent of input data norm, the trade-off parameter is accepted. If the fit is poor, I decrease the trade-off parameter by $1/10$. Conversely, if the algorithm overfits the data, I increase λ by a factor of 10.

3.5 Preconditioned implementation of RLSM

A negative aspect of RLSM is its computational cost. Each CG iteration requires one full migration and de-migration sequence, which clearly limits the feasibility of RLSM for industrial applications.

Preconditioning strategies for iterative solvers can help to speed up convergence and have been extensively studied in applied mathematics (Saad, 1991; Hanke and Hansen, 1993). Indeed, preconditioning schemes have been successfully applied in prestack imaging by coupled linearized inversion (Ehinger and Lailly, 1991), wave-equation least-squares migration (Prucha and Biondi, 2002), interpolation problems (Fomel and Claerbout, 2003) and Radon processing (Trad et al., 2003). In my implementation, equation 3.2 is solved with the addition of a simple change of variable:

$$\mathbf{z} = \mathbf{D}_{1h_x}\mathbf{m}\,. \tag{3.16}$$

The substitution of $\mathbf{m}$ in equation 3.2 leads to

$$\begin{aligned} F(\mathbf{z}) &= ||\mathbf{W}(\mathbf{d} - \mathbf{LPz})||^2 + \lambda^2||\mathbf{z}||^2 \\ &\equiv ||\mathbf{W}(\mathbf{d} - \tilde{\mathbf{L}}\mathbf{z})||^2 + \lambda^2||\mathbf{z}||^2, \end{aligned} \tag{3.17}$$

where $\mathbf{P}$, in theory, is the inverse of $\mathbf{D}_{1h_x}$. Here, rather than inverting $\mathbf{D}_{1h_x}$, I replace $\mathbf{P}$ by an operator that behaves similarly as the inverse of $\mathbf{D}_{1h_x}$. If $\mathbf{D}_{1h_x}$ is a discrete operator, we can think of it as a high-pass operator or filter. Therefore, $\mathbf{P}$ must be a low-pass operator. In my implementation, applying $\mathbf{P}$ is equivalent to applying a 1-D low-pass filter (Hamming window) to the image gathers. The convolution, in this context, is used to remove artifacts arising from incomplete sampling, additive noise in the original data and operator artifacts (Kuehl and Sacchi, 2003; Wang et al., 2004; Wang et al., 2005). The new cost function is similar but not identical to equation 3.2, since $\mathbf{P}$ is not exactly equivalent to the inverse of $\mathbf{D}_{1h_x}$. However, the solutions are expected to be of similar character and quality. Indeed, my tests confirm that both techniques yield almost indistinguishable results, but the preconditioned solution is reached significantly faster.

Mathematically, the logic behind this step is that a good preconditioner will change the distribution of eigenvalues of the operator $\mathbf{L}'\mathbf{L}$ (Saad, 1991). Proper preconditioning will introduce clustering of large eigenvalues and, consequently, the CG method will require fewer iterations to minimize the cost function F. Although I have no formal proof of this, I found that a low-pass filter as the preconditioner provides a good solution within only a few CG iterations of least-squares migration. My preconditioning strategy is similar to the "good-pass" operator (Ronen et al., 1995) for de-aliasing the dip moveout operator (DMO). Here, the "good-pass" operator is the Hamming smoothing window, and the "bad-pass" operator is the first-order derivative.

Furthermore, preconditioned least-squares migration (PLSM) allows us to set the trade-off parameter λ to zero and let the number of CG iterations control the data fitting (Hanke and Hansen, 1993). This saves the time otherwise required for finding a proper trade-off parameter.

3.6 A toy example: inversion of a smooth model

To evaluate the performance of the regularized least-squares inversion methods, I prepared a simple dataset by convolving a ricker wavelet with a sinc style model (see Figure 3.1). Five percent white noise is added to see if the algorithms can deal with inaccurate observations.

The first test is to monitor the inverted model at different CG iterations of the regularized least-square inversion, which is similar to previously described RLSM, but I replace the one-way wave-equation operator with a simple convolution matrix. The result is displayed in Figure 3.2. It is clear that with more CG iterations, the solution becomes

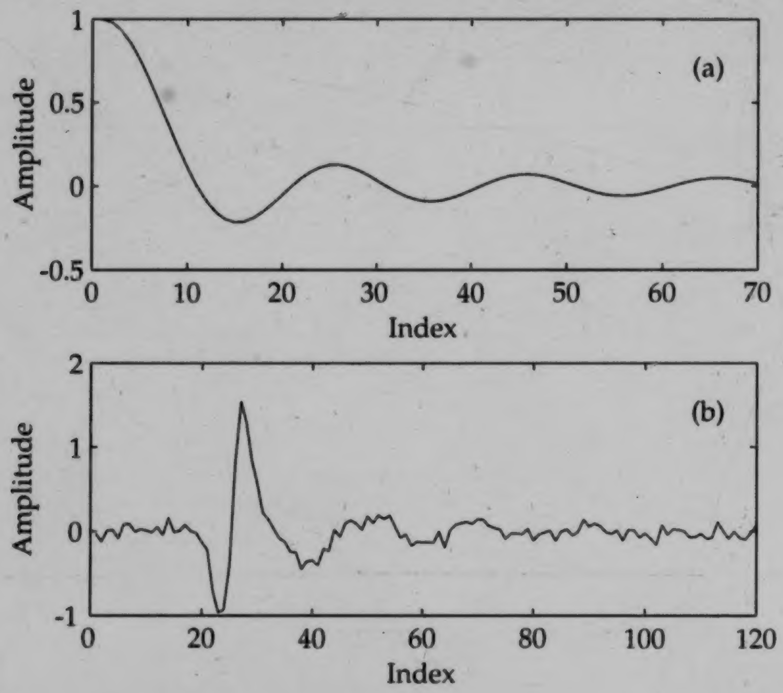


Figure 3.1: A smooth model and the simulated noisy data. (a) Model. (b) Noisy data prepared by convolving (a) with a Ricker wavelet and adding 5% white noise.

smoother and closer to the real model. After 30 iterations, a satisfactory result is obtained. Please note that a good trade-off parameter $\lambda = 1.0$ is found by the trial-and-error strategy.

The second test is to analyze the behavior of the algorithm with the number of preconditioned CG iterations. A Hamming filter with a length of 15 is used as the preconditioner. The inverted model behaves differently as that of the regularized least-squares inversion without preconditioning. As shown in Figure 3.3 With more CG iterations, the solution first improves, and then converges to a sub-optimal solution. The reason is that the trade-off parameter is set to zero, and the data fitting is controlled by CG iterations. With too many iterations, the algorithm will try to fit the noise. However, a few Preconditioned CG (PCG) iterations are sufficient to achieve a result similar to that of the convergent regularized CG algorithm. In this example, only four iterations of PCG are required to produce a good-matching result.

The advantage of PCG is evident in a comparison of the data misfit (see Figure 3.4). PCG algorithm fits the data much faster than the Regularized CG (RCG). Five iterations of PCG provides the same level of data fitting as 30 iterations of RCG. Usually, the algorithm is stop when the misfit curve becomes flat. The normalized eigenvalues of $\tilde{\mathbf{L}}'\tilde{\mathbf{L}}$ for the two methods are compared in Figure 3.5, which confirms that preconditioning has caused clustering of eigenvalues.

The third test is to examine the influence of the Hamming filter length on the solution of the preconditioned CG. Three filters ($NF = 4$, $NF = 15$ and $NF = 30$) were used to analyze the behavior of the algorithm. Four PCG iterations are completed for each test. The solutions are displayed in Figure 3.6. It can be seen that with the increase of filter length, the accuracy first increases, and then decreases. Empirically, a filter with a length from one quarter to one third of the model length provides a satisfactory solution.

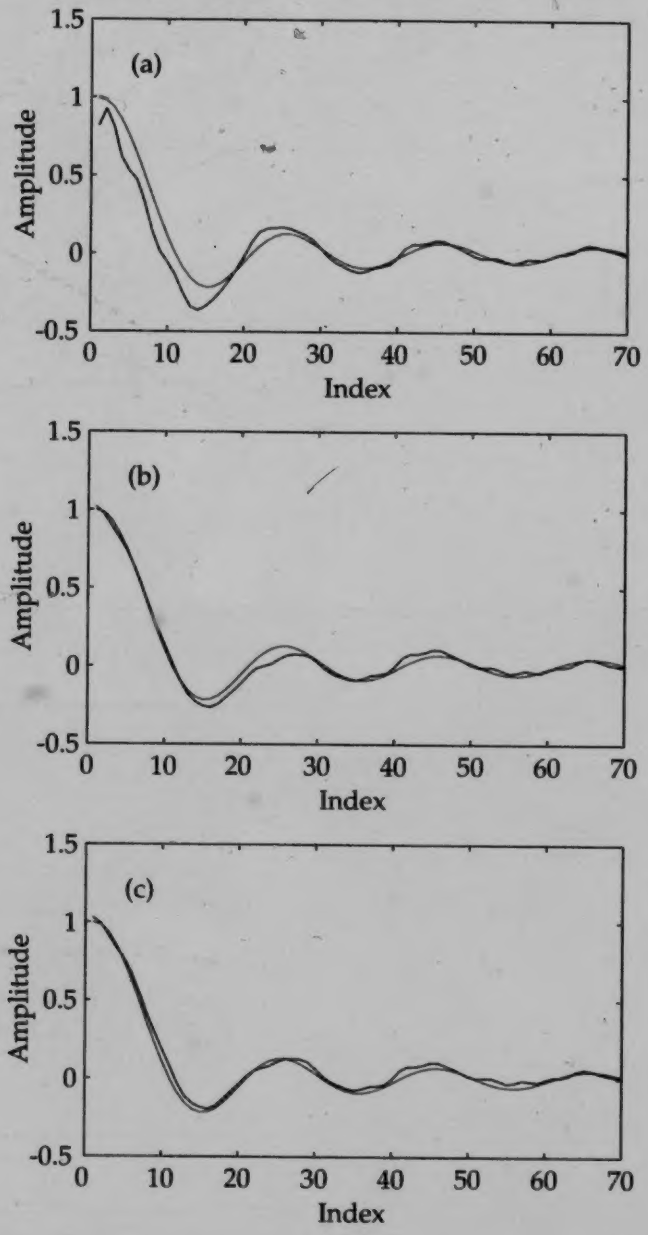


Figure 3.2: Reconstructed model (Blue) vs. the true model (Red) at different regularized CG iterations. (a) The fourth iteration. (b) The 15th iteration. (c) The 30th iteration.

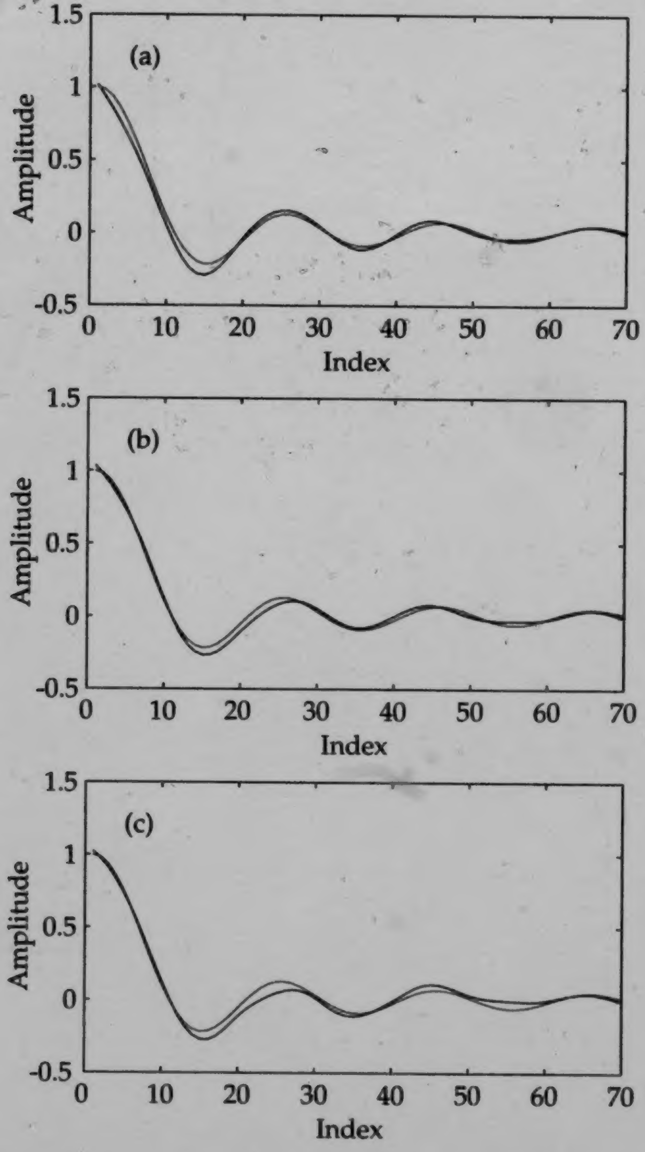


Figure 3.3: Reconstructed model (Blue) vs. the true model (Red) at different preconditioned CG iterations. (a) The second iteration. (b) The fourth iteration. (c) The eighth iteration.

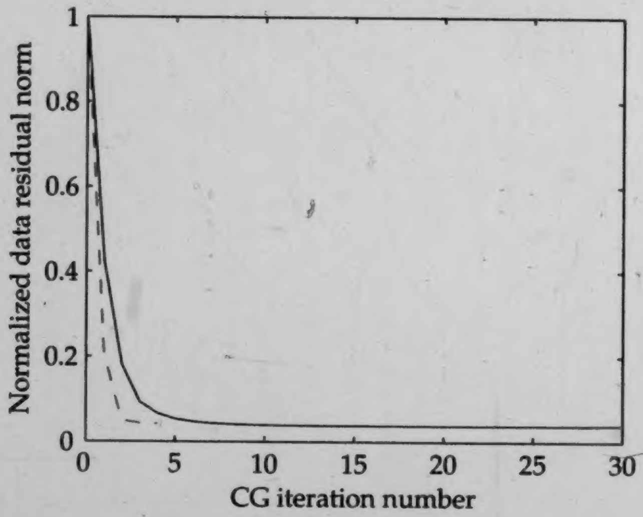


Figure 3.4: Data misfit comparison between the regularized CG (Blue, $\lambda = 1.0$) and the preconditioned CG (Red dashed, $NF = 13$).

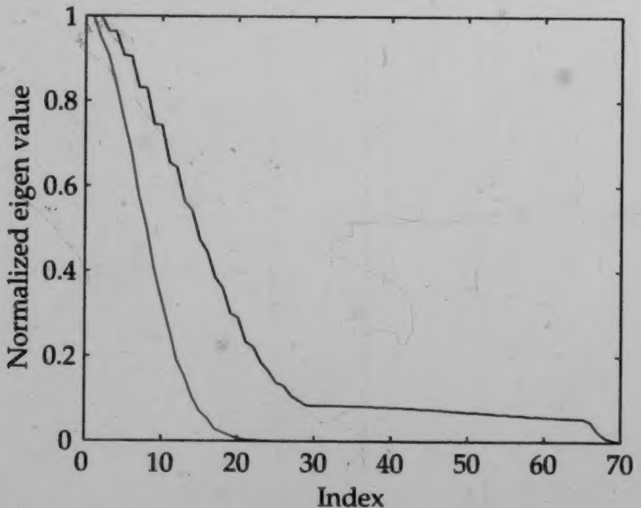


Figure 3.5: Eigenvalue comparison between the regularized CG (Blue, $\lambda = 1.0$) and the preconditioned CG (Red dashed, $NF = 13$)

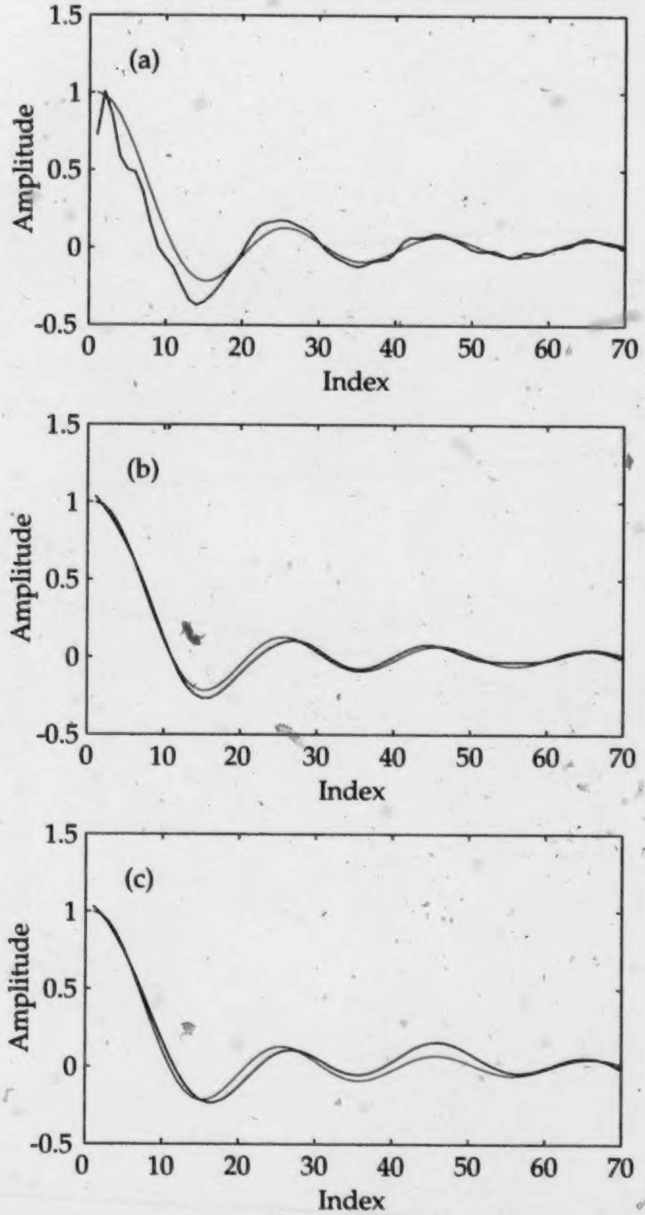


Figure 3.6: The reconstructed model (Blue) vs. the true model (Red) by the preconditioned CG iterations using different filter length (NF). (a) $NF = 3$. (b) $NF = 1$. (c) $NF = 31$.

3.7 Sparse least-squares migration

One possible way to further enhance the resolution and attenuate artifacts is by taking advantage of the solution itself. Iteratively using the result as a model-space regularization can lead to high-resolution artifact-free seismic images. This idea has been used in many fields of signal and image processing (Sacchi and Ulrych, 1995; Charbonnier et al., 1997; Youzwishen, 2001; Sacchi et al., 2003; Trad et al., 2003; Downton and Lines, 2004). In this thesis, I combine a model-dependent sparse regularization and a model-independent smoothing regularization to estimate AVP common image gathers. Model-dependent sparse regularization is introduced via a non-quadratic norm (Cauchy norm). Smoothing is implemented via a convolutional operator applied to AVP common image gathers along the ray parameter direction. This idea is used to develop an algorithm to simultaneously improve the structural interpretability and amplitude accuracy of seismic images.

It is important to point out the similarities between this algorithm and methods for impedance inversion based on sparse spike deconvolution of post-stack cubes (Oldenburg et al., 1983; Debeye and van Riel, 1990). In principle, I use very similar concepts to find a solution that exhibits pre-defined properties such as sparseness, smoothness,etc. The main difference of the proposed method with respect to sparse spike inversion strategies is that the operator is a one-way forward modelling operator rather than a convolutional kernel. In addition, the inversion results are in depth and the input data are prestack volumes as opposed to time-domain reflectivity estimates and post stack volumes, respectively. I believe that the proposed method provides a unifying thread between convolution-based sparse spike inversion and regularized migration/inversion methods.

Combining the smooth regularization in the ray parameter direction and the sparse regularization in the depth direction, the following new cost function is proposed:

$$F(\mathbf{m}) = ||\mathbf{W}(\mathbf{Lm} - \mathbf{d})||_2^2 + \lambda^2 R(\mathbf{SH}(\mathbf{m})), \tag{3.18}$$

where $\mathbf{H}$ is a high-pass filter to penalize rapid variations along the ray parameter axis, $\mathbf{S}$ is a stacking operator that converts common image gathers to a stacked image, and R is a model-dependent functional (Cauchy norm) used to enforce sparseness.

By adopting the aforementioned preconditioning strategy (Prucha and Biondi, 2002; Wang et al., 2004), the cost function can be expressed as follows

$$F = ||\mathbf{W}(\mathbf{LPz} - \mathbf{d})||_2^2 + \lambda^2 R(\mathbf{Sz}), \tag{3.19}$$

where $\mathbf{P}$ is the preconditioning matrix, and $\mathbf{z}$ is the model modified by the preconditioner. It is clear that the final solution is $\mathbf{m} = \mathbf{Pz}$. The problem can be efficiently solved by Iterative Re-weighted Least-squares (IRLS) (Scales and Smith, 1994). The cost function at the k-th iteration of the IRLS algorithm is given by

$$F(\mathbf{z}_k) = ||\mathbf{W}(\mathbf{LPz}_k - \mathbf{d})||_2^2 + \mu^2||\sqrt{Q_{k-1}}\mathbf{Sz}_k||_2^2, \tag{3.20}$$

where Q_{k-1} is a diagonal weighting matrix with diagonal elements given by

$$Q_{ii}^{k-1} = \frac{1}{1 + (m_{si}^{k-1}/\sigma_{ms}^{k-1})^2}. \tag{3.21}$$

In the above expression, m_{si}^{k-1} is the i-th element of the vector $\mathbf{Sz}$ at the $(k-1)$-th iteration of IRLS. Finally, σ_{ms}^{k-1} is a scale parameter, which is empirically set to some percentage of the maximum amplitude of the aforementioned vector $\mathbf{Sz}$. Application of the IRLS method involves properly choosing two hyperparameters, μ and σ_{ms}^{k-1}. The latter can be reduced to selecting one hyperparameter, δ, by using the following expression: $\sigma_{ms}^{k-1} = \delta \cdot max(|\mathbf{m}_s^{i-1}|)$, where $\mathbf{m}_s = \mathbf{Sz}$.

The IRLS algorithm with two regularizations is summarized as follows:

- Initialize $\mathbf{z}$ with zeros and compute $\mathbf{Q}$
- Minimize cost function 3.20 via the CG algorithm.
- Update the diagonal matrix $\mathbf{Q}$, and restart the CG algorithm.

The above procedure requires about 3-4 updates (iterations) to obtain a solution that is sparse in depth and smooth with respect to the ray parameter.

It can be seen that the first iteration of IRLS algorithm is identical to PLSM with a trade-off parameter control. The solution is prewhitened, and therefore, it is more stable than the previously described PLSM.

The two hyperparameters, μ and δ, can be related to each other. Consider two pairs of hyperparameters μ, δ and μ', δ' such that

$$\begin{aligned} \mu' &= c\mu, \\ \delta' &= \frac{1}{c}\delta, \end{aligned} \tag{3.22}$$

where c is a constant. The i th element of the diagonal weighting matrix for the second pair of hyperparameters can be derived as bellow

$$\begin{aligned} Q'_{ii} &= \frac{1}{1+(m_{si}/(m_M \cdot \delta'))^2} \\ &= \frac{1}{1+(m_{si}/(m_M \cdot \delta))^2 \cdot c^2} \\ &\approx \frac{1}{c^2}\frac{1}{1+(m_{si}/(m_M \cdot \delta)^2)} \\ &= \frac{1}{c^2}Q_{ii} \end{aligned} \tag{3.23}$$

where $m_M = max(|\mathbf{m}_s|)$. The approximation is achieved by using the fact that the term $(m_{si}/(m_M \cdot \delta)^2)$ dominates the denominator for non-zero m_{si}. Therefore,

$$\begin{aligned} \mu'^2 \|\sqrt{\mathbf{Q}'_{k-1}} \cdot \mathbf{S}\mathbf{z}_k\|_2^2 &\approx c^2 \cdot \mu^2 \cdot \frac{1}{c^2} \|\sqrt{\mathbf{Q}_{k-1}} \cdot \mathbf{S}\mathbf{Z}_k\|_2^2 \\ &= \mu^2 \|\sqrt{\mathbf{Q}_{k-1}} \cdot \mathbf{S}\mathbf{z}_k\|_2^2, \end{aligned} \tag{3.24}$$

The regularization norm is calculated from non-zero image points since the contribution of zero image points is zero. Equation 3.24 shows that pairs of δ and μ with a constant product lead to similar solutions. However the tests on sparse deconvolution further bellow show that too large values of δ yield low-resolution results. Therefore in practice, δ is set to a small value, for example, 0.02, whereas, the the other hyperparameter μ is adjusted to obtain a satisfactory fitting.

3.8 Hyperparameter tests

To examine the influence of the hyperparameters on the inversion result, I played with a sparse deconvolution problem by varying the parameters. The data were prepared by convolving a sparse reflectivity model with a Ricker wavelet and adding some white noise (see Figure 3.7).

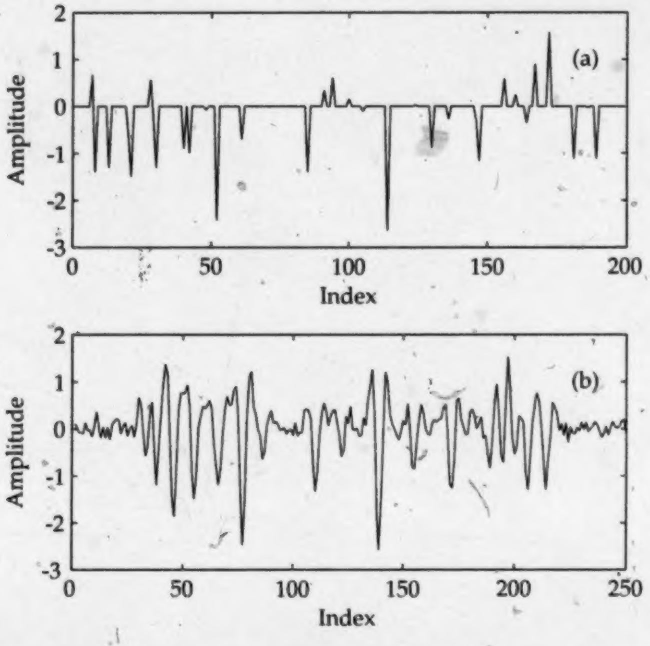


Figure 3.7: Sparse reflectivity model and the simulated noisy seismic data. (a) The reflectivity model. (b) The noisy data prepared by convolving (a) with a Ricker wavelet and adding 5% white noise.

The first test is to fix the trade-off parameter μ and change the scaling parameter δ. As shown in Figure 3.8, a too small scaling parameter may cause instability of the algorithm. Many oscillatory artifacts are present in the solution. With the increase of the scaling parameter, the solution becomes cleaner, but a too large value will also suppress tiny detail of the model. This is a typical trade-off situation.

The second test is to fix the scaling parameter δ and only vary the trade-off parameter μ. Similar phenomenon is observed in Figure 3.9. A proper trade-off parameter should be found to avoid instability or over regularization.

The third test is to fix the product of two hyperparameters. The result (see Figure 3.10) confirms the previously proved rule about the relationship between the trade-off and scaling parameters. However, the scaling parameter can not be too large. As shown in Figure 3.10e, resolution is lost with a large scaling parameter. In practice, usually we

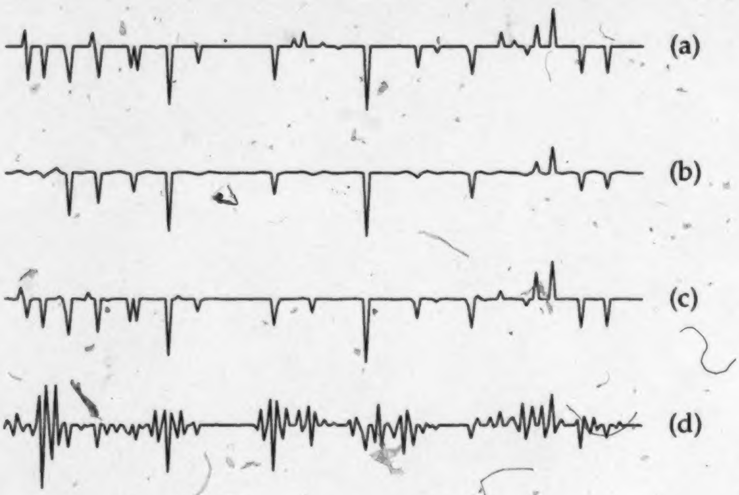


Figure 3.8: True reflectivity model vs. inverted models by using a constant trade-off parameter ($\mu = 4.0$). (a) The true model. (b) $\delta = 0.2$. (c) $\delta = 0.02$. (d) $\delta = 0.002$.

can fix the scaling parameter to some percentage, for example, 0.02, and adjust the trade-off parameter to fit the data.

3.9 A simple example: multi-channel deconvolution

I tested the idea of forcing sparseness and smoothness at the same time by a simple multi-channel deconvolution problem. This is an unrealistic scenario in seismic deconvolution but yet it is a good example to test the algorithm. The procedure, however, could be used to deconvolve time-migrated common image gathers. Figure 1a is a time domain model with 20 offsets. I convolved the model with a zero-phase wavelet, and removed three offsets (trace number equals 2, 6 and 9) to also test the procedure in situations of missing information. The data are portrayed in Figure 3.11b.

I compared two methods of inversion, preconditioned LS inversion and sparse LS inversion. Figure 1c is the result of the preconditioned LS inversion after 50 iterations of the CG algorithm. It is evident that the inversion successfully fill in the gap in the incomplete data. However, the vertical resolution is not satisfactory. Spurious sidelobes

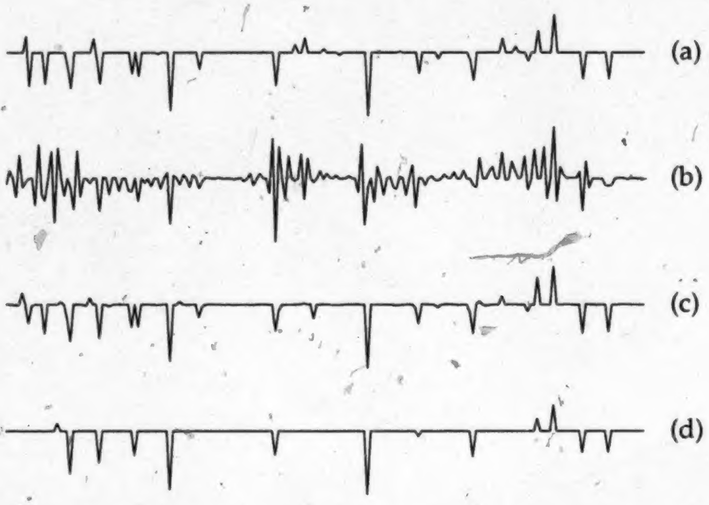


Figure 3.9: True reflectivity model vs. inverted models by using a constant scaling parameter ($\delta = 0.02$). (a) The true model. (b) $\mu = 0.4$. (c) $\mu = 4.0$. (d) $\mu = 40$.

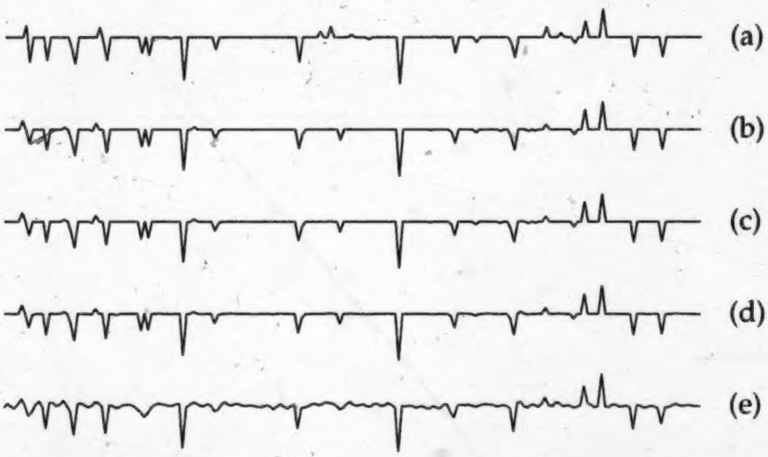


Figure 3.10: True reflectivity model vs. inverted models by different sets of hyperparameters with a constant product. (a) The true model. (b) $\mu = 400, \delta = 0.0002$. (c) $\mu = 40, \delta = 0.002$. (d) $\mu = 4, \delta = 0.02$. (e) $\mu = 0.4, \delta = 0.2$

are present in the inverted reflectivity model. On the other hand, the sparse inversion provides a superior result (Figure 3.11d). It is almost identical to the real reflectivity model. The wavelet is properly compressed. In addition, the AVO signature is preserved.

3.10 True amplitude migration

In the last few sections, I have described a few iterative inversion methods to reconstruct angle dependent reflectivity. There exist more economical ways to approximately solve the inverse problem. For example, if seismic data are well sampled and the velocity structure is not very complex, an approximate solution can be acquired by applying correction weights in the downward continued wavefield to preserve the amplitudes. The correction aims at removing the effects introduced by the imaging Jacobian as a consequence of the imaging procedure. Mathematically, the correction estimates the inverse of the smearing operator $L^T L$. This idea was first introduced by Stolt and Benson (1986) to address the conversion of integral variables during migration. The theory was further developed and applied to angle-domain and ray parameter domain imaging methods (Sava et al., 2001a). As shown bellow, the imaging Jacobian can also be explained as a scaling factor of the forward and inverse Fourier transforms involved in the modelling and migration procedures (Biondi, 2003a; Sava et al., 2001b).

The imaging condition of the wave continuation method can be expressed as:

$$\tilde{R}(z) = \int_{-\infty}^{+\infty} d\omega e^{-ik_z z} \tilde{P}(\omega), \tag{3.25}$$

where $\tilde{R}$ is the estimated local reflectivity wavefield, and $\tilde{P}$ is the recorded seismic response in the surface, both of which are expressed in the Fourier domain. As previously discussed, the seismic data can be considered as the summation of propagated waves from all depth, i.e.

$$\tilde{P}(\omega) = \int_{-\infty}^{+\infty} dz e^{ik_z z} R(z), \tag{3.26}$$

where $R(z)$ is the true local reflectivity wavefield in the Fourier domain. Please note that the integral range is extended to negative infinity by assuming the reflectivities are zeros

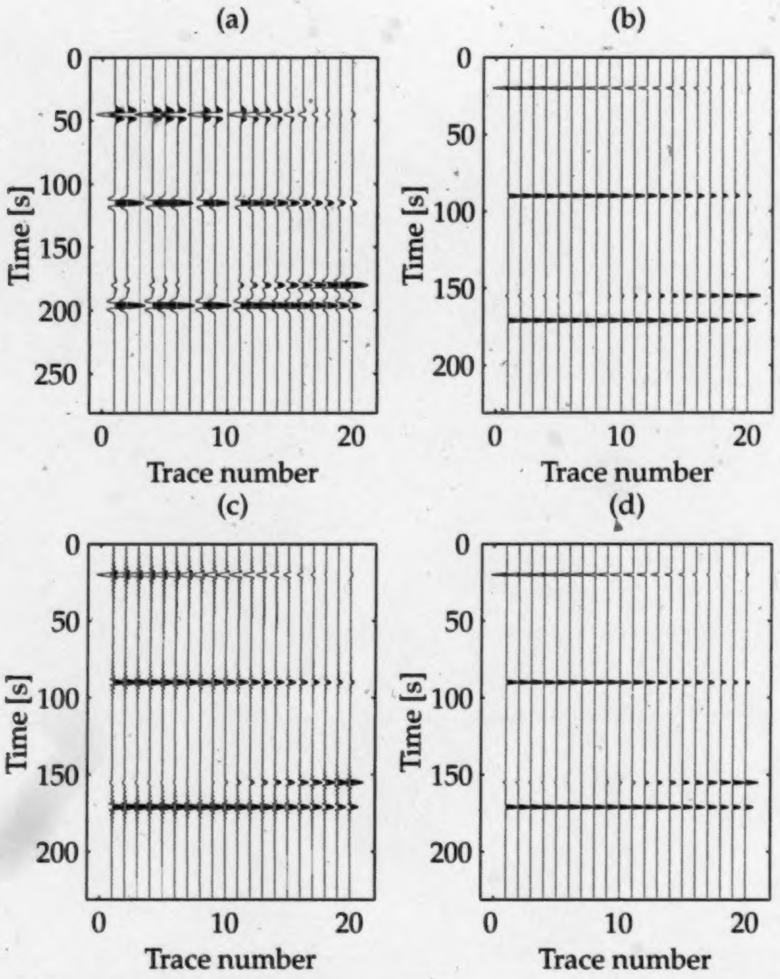


Figure 3.11: A multi-channel deconvolution example to compare linear inversion and non-linear inversion. (a) Incomplete multi-channel data. (b) Reflectivity model. (c) Linearly inverted model. (d) non-linearly inverted model (sparse solution)

above the surface. This is reasonable since above the surface there is just air, and the seismic wave dies away quickly in the air. Combining the above two equations leads to

$$\tilde{R}(z) = \int_{-\infty}^{+\infty} d\omega e^{-ik_z z} \int_{-\infty}^{+\infty} dz' e^{ik_z z} R(z'), \tag{3.27}$$

The physical meaning of this equation is clear if we change the integral variable from the temporal frequency to the vertical wavenumber:

$$\begin{aligned} \tilde{R}(z) &= \int_{-\infty}^{+\infty} dk_z e^{-ik_z z} \frac{d\omega}{dk_z} \int_{-\infty}^{+\infty} dz' e^{ik_z z} R(z') \\ &= \frac{d\omega}{dk_z} \int_{-\infty}^{+\infty} dk_z e^{-ik_z z} \int_{-\infty}^{+\infty} dz' e^{ik_z z} R(z') \\ &= \frac{d\omega}{dk_z} R(z), \end{aligned} \tag{3.28}$$

by recognizing the chaining of the forward and inverse Fourier transforms in the manipulation. This equation tells us that the reflectivity model calculated by the adjoint operator or migration is a non-constantly scaled version of the true reflectivity model. Therefore, amplitude fidelity can be acquired by removing such scaling factor. To preserve amplitude during DSR wave-equation migration, the downward continued wavefield should be divided by the imaging Jacobian, $d\omega/dk_z$.

The formula of the imaging Jacobian for true amplitude migration can be derived right away given the expression of the vertical wavenumber. The derivation entails calculating the derivative of the vertical wavenumber over the temporal frequency. However, the solution can be different for different imaging conditions. For example, angle-domain imaging condition (Wapenaar and Berkhout, 1986b; Wapenaar et al., 1999; Sava et al., 2001a) maps common image gathers (CIGs) from the local wavefield in terms of offset wavenumbers. Therefore the derivative is calculated by assuming constant offset wavenumbers. Similarly the imaging Jacobian can be derived by assuming constant offset ray parameter for ray parameter domain imaging methods (Prucha et al., 1999). Appendix C provides the imaging Jacobian formulas for both methods.

Compared with the conventional migration, the described true amplitude migration provides more accurate result and requires only a little extra computational effort. Therefore, the method provides an attractive option for calculating models that may be suitable for amplitude analysis. However two reasons hamper the wide application of this method. First, the method is derived by assuming constant velocity. Thus it can not properly handle the situation when there exist velocity variations. Second, the amplitude correction does nothing to suppress the effects of data missing and noise, which is a very common challenge for real seismic data processing.

3.11 Summary

In this chapter, I analyzed regularized least-squares migration. Both quadratic and non-quadratic regularizations are formulated and tested via a simple deconvolution problem. The acquired experience on parameter selection is especially valuable since migration is a much more expensive procedure, and a good parameter-selection strategy saves trial-and-error time. Important amount of work, especially in the statistics literature, has been done to deal with the problem of automatic parameter selection (Akaike, 1974). Unfortunately, this research often deals with the solution of small parameter fitting problems, as the ones arising in time series analysis.

In the next two chapters, I will test the methods on migration problems with synthetic and real data to further validate the application of regularized migration.

Chapter 4

Synthetic data examples

In this chapter, I test the theory of wave-equation migration/inversion with synthetic data examples including the 2-D zero-offset SEG/EAGE data, the 3-D SEG/EAGE narrow-azimuth data, the 2-D prestack Marmousi data, and the 3-D common-azimuth data of a flat model and 2-D data of a wedge model. The quality of structural imaging capability and amplitude fidelity of various imaging methods is extensively examined.

4.1 Migration

SEG/EAGE salt body model

The efficiency for dealing with complex geological structure and velocity variations is an important criterion at the time of evaluating an imaging algorithm. For this reason, it is important to produce some benchmark data so that any developed imaging methods can be fairly evaluated and compared. With the joint effort of many geologists, geophysicists and various organizations, the 3-D SEG/EAGE salt body data were created in the 90's. The model simulates a typical geological setting in the U.S. Gulf of Mexico that contains a big salt dome surrounded by a balanced cross section (see Figure 4.1, Aminizadeh et al., (1994)). The sharp velocity contrast between the salt dome and the background rock poses a challenge to seismic imaging since most energy is reflected back to the surface. It is especially difficult to illuminate the area beneath the salt body. Furthermore numerous faults and thrusts were included to simulate a complicated situation. An ideal imaging method should be able to address these two issues.

For 2-D migration test, a zero-offset data were produced based on the 'exploding re-

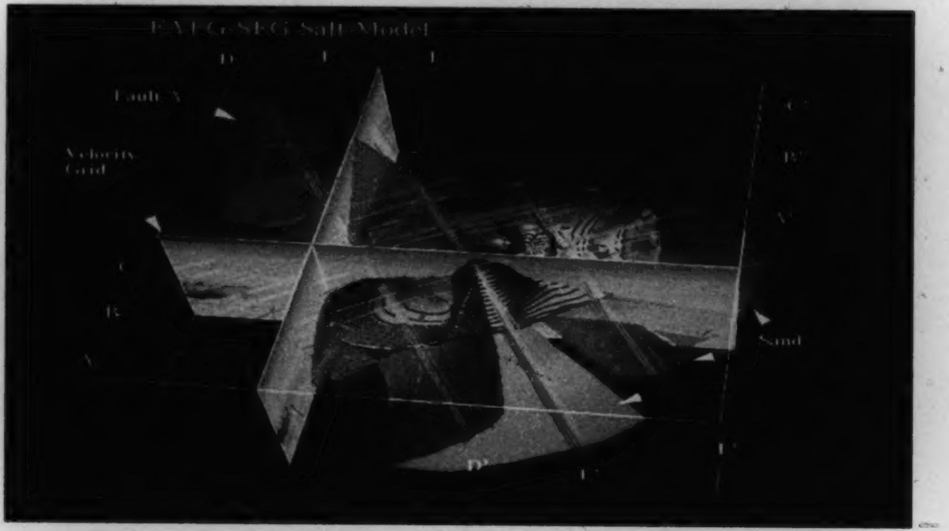


Figure 4.1: 3-D SEG/EAGE salt model (Aminzadeh et al., 1994). The model embodies typical complex gulf structures like salt domes, faults and sand bodies and lenses. The dimensions of the model are 13.5 × 13.5 × 4.2 km in the x, y, z directions, respectively. The 2-D velocity profile AA′ is shown in Figure 2.3, which is used to synthesize the 'exploding reflector' zero-offset dataset

flector' theory for the profile AA′. The relative data have been shown in Chapter 2 and used to test the Kirchhoff migration method. In this chapter, this dataset is used to compare velocity correction techniques including split-step, split-step PSPI, PSPI and FFD (Fourier Finite Difference), etc. As shown before, 2-D or 3-D DSR migration actually involves calculating two square roots. The velocity correction for 2-D/3-D prestack migration can be regarded as two single-square-root corrections. Therefore the quality comparison for 2-D poststack cases also hold for 3-D cases though 3-D is more complicated with a larger space of velocity candidates.

The main frame of the geological setting of 3-D SEG/EAGE salt body model is made up of layers with gradually changing velocities and some spiky velocity perturbations (for example, see Figure 2.3). The density is set to a constant everywhere due to the limitation of the adopted finite difference modelling method . Therefore, the data are not suitable for amplitude studies (no distinct AVA phenomena), but excellent for structural analysis.

2-D SEG/EAGE poststack data

The 2-D SEG/EAGE poststack data contain 1290 zero-offset seismic traces. The sample interval is 0.008 second, and each trace has 626 samples. The peak frequency is $15Hz$. For accuracy, the migration frequency range is picked between 0 and $55Hz$. Figure 4.2 is the result provided by the split-step zero-offset migration algorithm. The method successfully resolves the faults above the salt body and salt flank with high dip angles. However, the sub-salt structure is not properly placed. Obviously, the flat layer at depth $z = 3.6$ km is distorted under the salt dome. The discontinuity at the distance $x = 7$ km could be mistakenly explained as a fault. Besides, the fault system under the salt dome is disturbed by many high frequency artifacts. On the other hand these artifacts are cleaned up by interpolation methods like split-step PSPI (Figure 4.3) and PSPI (Figure 4.4). The latter two methods provide very similar result. Both methods accurately image the faults under the salt dome. However, the dipping salt flank on the left is not preserved so well as the simple split-step method.

The result of FFD is shown in Figure 4.5. Since higher-order approximation is adopted in the algorithm, the high angle dipping events are better preserved. However, some spurious artifacts are present throughout the whole image. This is a drawback of the algorithm. Furthermore, the involved finite difference method is difficult to implement for iterative inversion. Therefore, the method is not used for inversion in this thesis.

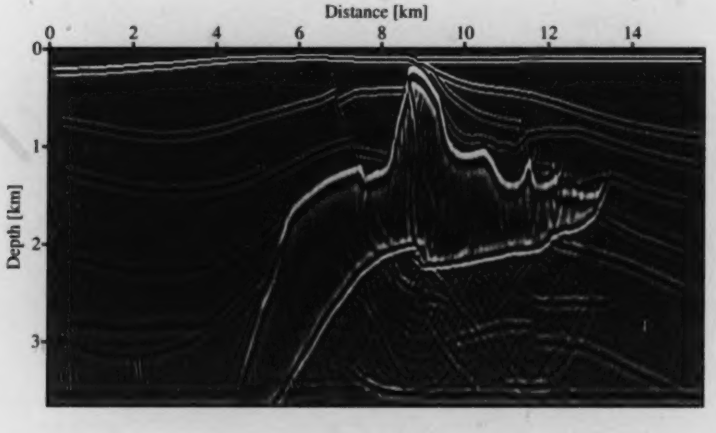


Figure 4.2: Migrated image of the 'exploding-reflector' data by split-step zero-offset migration.

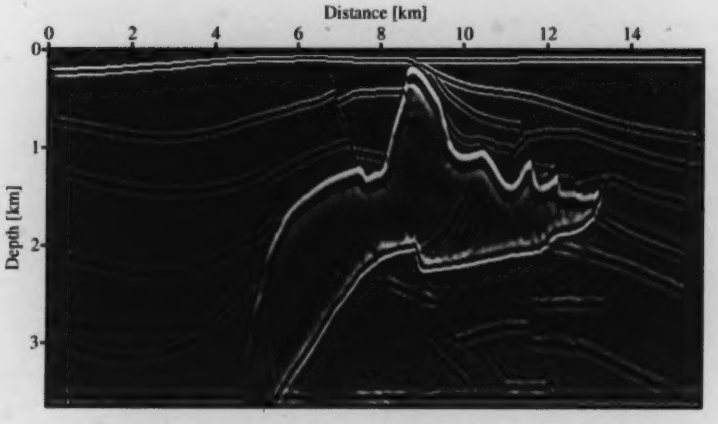


Figure 4.3: Migrated image of the 'exploding-reflector' data by split-step PSPI zero-offset migration.

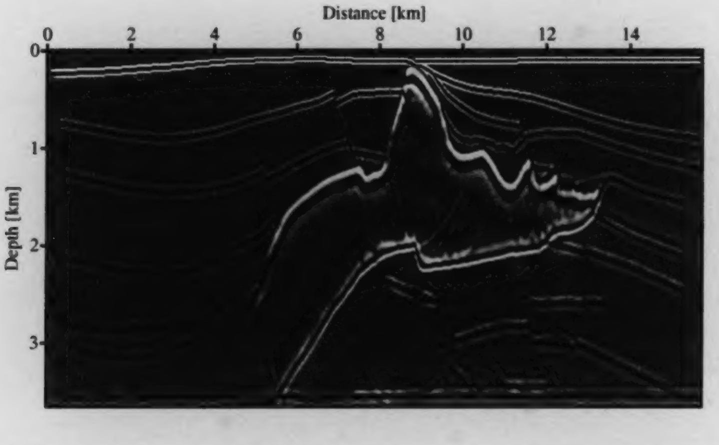


Figure 4.4: Migrated image of the 'exploding-reflector' data by PSPI zero-offset migration.

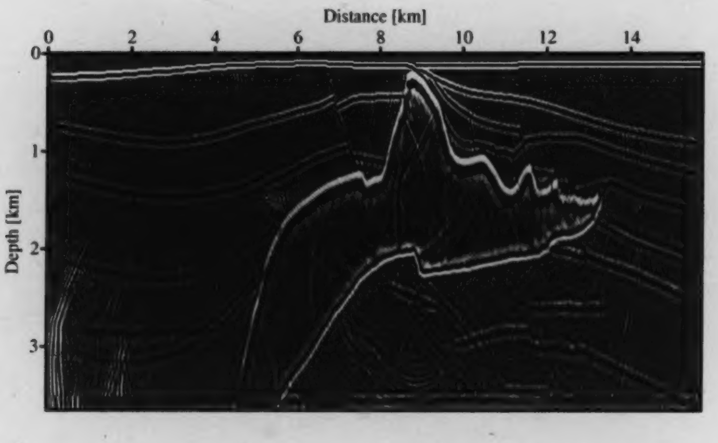


Figure 4.5: Migrated image of the 'exploding-reflector' data by Fourier Finite Difference zero-offset migration.

3-D SEG/EAGE narrow-azimuth data

Narrow-azimuth marine data are usually acquired at the surface of an ocean with floating geophones that are attached to parallel streamers (see Figure 4.6). A boat tugs the streamer back and forth in the survey area to scan the geological structure under the ocean bottom. This geometry allows us to well sample the seismic response in the inline direction and record that in the crossline direction with a much narrower aperture.

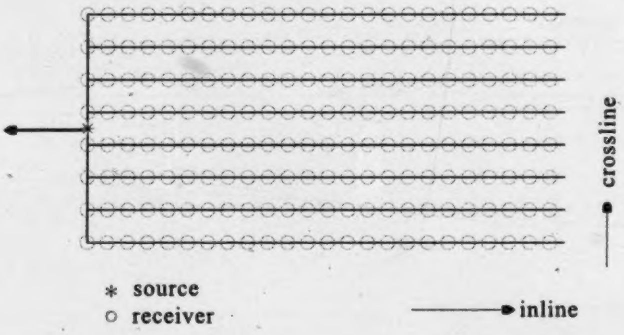


Figure 4.6: The geometry of marine data acquisition for 3-D SEG/EAGE salt model. Eight streamers are tugged by a tugboat (not shown in the figure). Each steamers contains 68 receivers in a spacing of 40 m..

The classic 3-D SEG/EAGE narrow-azimuth data (Phase C) were synthesized by using a typical marine acquisition geometry (Aminzadeh et al., 1994). The survey covered an area of 8 km ×8 km (see Figure 4.7). There are 50 shot lines with 160 m crossline spacing. Within each shot line, there are 95 shots every 80 m. Each streamer contains 68 receivers with 40 m spacing. The distance between two neighboring streams is 40 m. The boat sailed in the same direction for each shot line. With this survey design, the binned data have 17 folds and a half offset of 80 m. The inline and crossline intervals are both 20 m.

The data were calculated by a 3-D acoustic finite-difference modelling method and resampled with an interval of $0.008s$. To test the 3-D common-azimuth migration, I binned the data into a common azimuth (y direction). This simple treatment surely introduces some errors due to the trace rotation. However, the survey was conducted with a narrow-azimuth geometry, the errors should be minor. Despite this concern, the data provide good basis for testing the algorithm with the related complex geological structure. Furthermore, for the previously mentioned reason (spiky velocity perturbation), I focus on the structure instead of amplitude in this case. To alleviate aliasing, I interpolated the inline offsets so that the half offset spacing is 40 m.

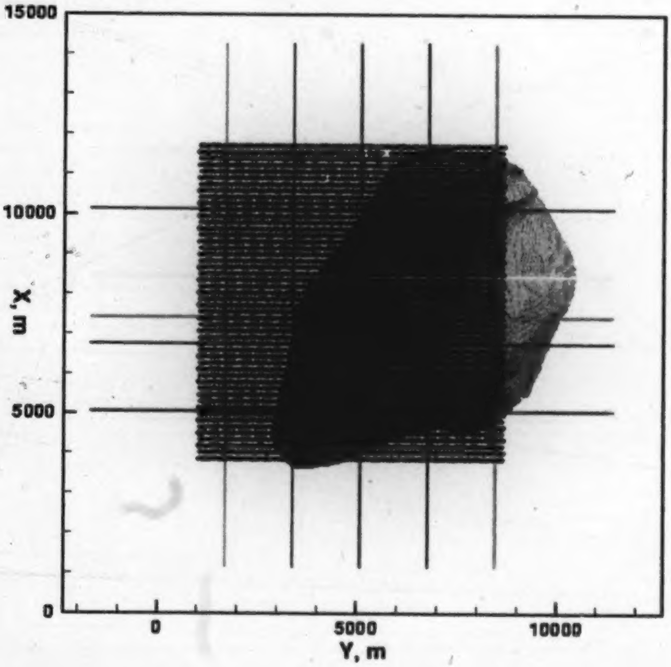


Figure 4.7: The survey area of 3-D SEG/EAGE Narrow azimuth data for the salt model (online figure provided by Sandia National Laboratories (1998)). x: crossline direction. y: inline direction. The asteroids symbolize the shot positions. The salt model is displayed in red as a reference. The longer color lines (5 inlines and 5 crosslines) show the location, $x = 5060$ m, 6760 m, 7400 m, 8440 m, 10140m, and $y = 1680$ m, 3380 m, 5060 m, 6760 m, 8440 m, where seismic images were extracted by Sandia National Laboratories.

Figure 4.8b and d show the stacked images at inline $x = 5460$ m, obtained with 2-D and 3-D common-azimuth migration respectively. It is obvious that the 3-D image is cleaner than the 2-D image. The structural accuracy is confirmed by checking with the velocity profile (see Figure 4.8a). The 2-D imaging method can not accurately place the normal fault on the right. Even the top of the salt body is distorted. The bottom of the salt body at depth $z = 3000$ m is imaged at a shallower depth.

Figure 4.8c and e are the common image gathers at $x = 5460$ m, $y = 6000$ m produced by 2-D and 3-D common-azimuth migration, respectively. The 3-D imaging method provides more coherent information in the ray parameter domain. Especially the event at depth $z = 1000$ m is less disturbed by the imaging artifacts in the 3-D result. The shallow events are obviously aliased in the 2-D result. On the contrary, 3-D result is much less affected. Under the salt body, some spurious events are present in the result of 2-D migration, while the true event at depth $z = 3500$ m stands out in the 3-D common image gather.

The Sandia National Laboratories (SNL) (1998) conducted a benchmark test on the narrow-azimuth and wide-azimuth data of the SEG/EAGE salt model using 3-D shot record migration (F-X (frequency-space) domain migration). They argued that narrow-azimuth acquisition does not provide more significant information than wide-azimuth acquisition for the SEG/EAGE salt model. If this is true, my comparison bellow shows that the common-azimuth migration improves over the benchmark short-record migration. In the following, I compare the migrated images by 3-D common-azimuth migration, applied to the narrow-azimuth data, and by 3-D shot record migration, applied to the wide-azimuth data. Note that the online benchmark results provided by SNL (1998) were calculated using a larger imaging aperture. For convenience of comparison, I have clipped the major part of the images.

Figure 4.9 compares the stacked images of inline $x = 6760$ m calculated by 3-D common-azimuth migration and shot migration. The common-azimuth migration is obviously better focused. In particular, the subsalt area is better illuminated and less affected by spurious imaging artifacts. The method provides a sharper shape of the shallow faults.

Figure 4.10 compares the stacked images at a different location ($x = 8440$ m), obtained with the two 3-D imaging methods. Note that the salt edge is not clear in the solution of 3-D shot migration (see Figure 4.10c). On the other hand, the salt flank in the middle is better resolved by the common-azimuth migration (see Figure 4.10b). Furthermore, we can see that common-azimuth migration improves the shallow fault image extending rom $y = 1000$ m to $y = 2000$ m.

Figure 4.11 shows the stacked images of a crossline ($y = 3380$ m). The result of common-azimuth migration (Figure 4.11b) is much cleaner than the 3-D shot migration. Even some of the weak events related to layers with small velocity contrast are resolved. The top of the salt body is more accurately placed, in contrast to the result of shot migration, which underestimates the size of the salt body by cutting the salt top.

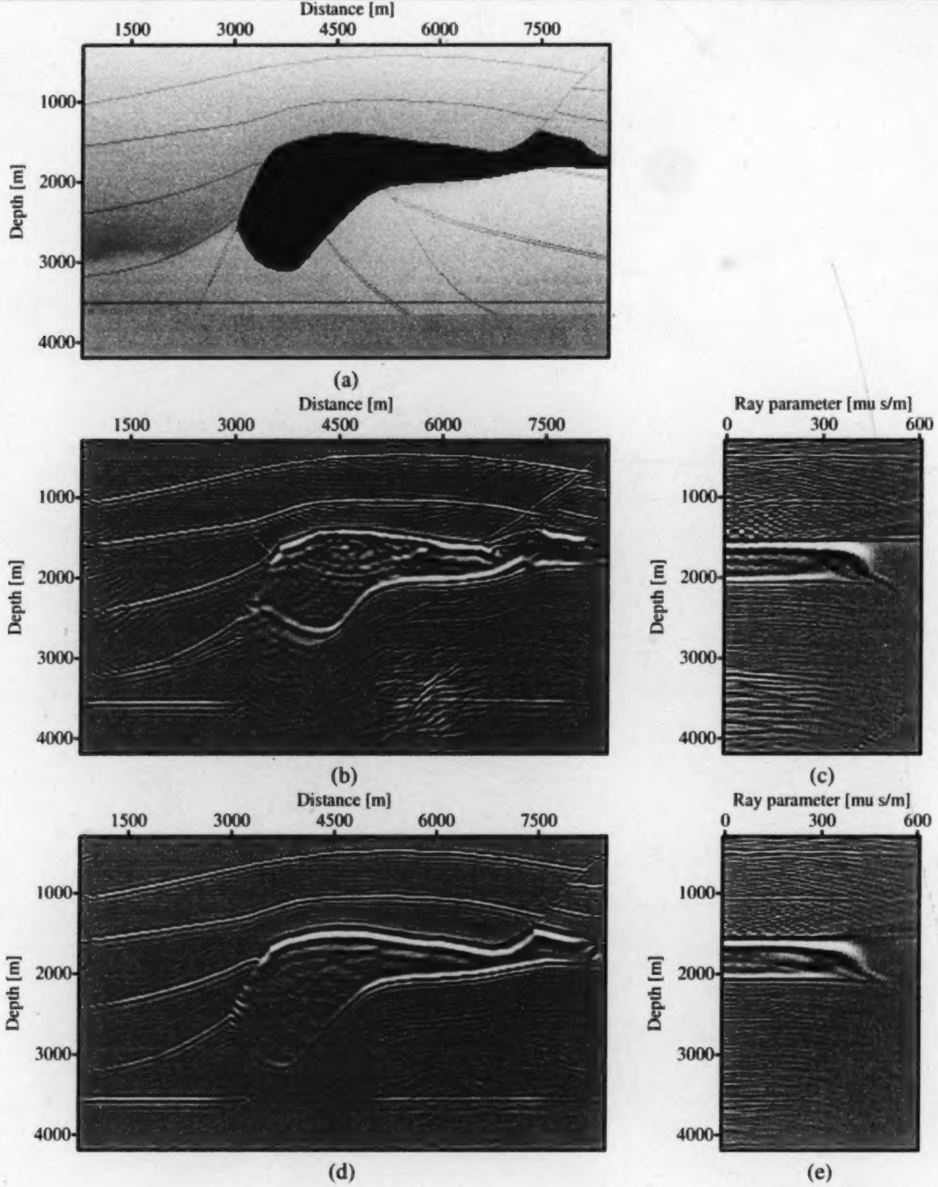


Figure 4.8: Image comparison between 2-D and 3-D migration at $x = 5460$ m. (a) velocity model. (b) Stacked image with 2-D migration. (c) Common image gather (CIG) with 2-D migration at $x = 5460$ m, $y = 6000$ m. (d) Stacked image with 3-D common-azimuth migration. (e) CIG with 3-D common-azimuth migration. Both methods use the split-step PSPI velocity correction.

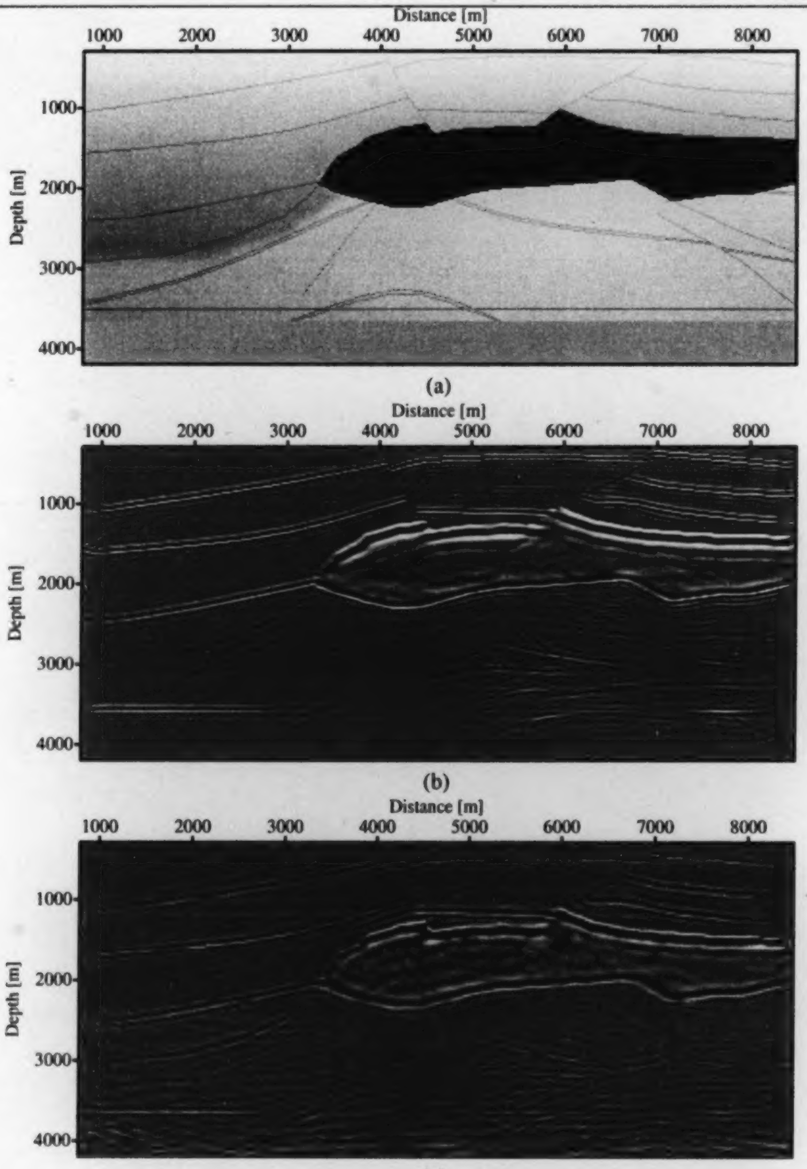
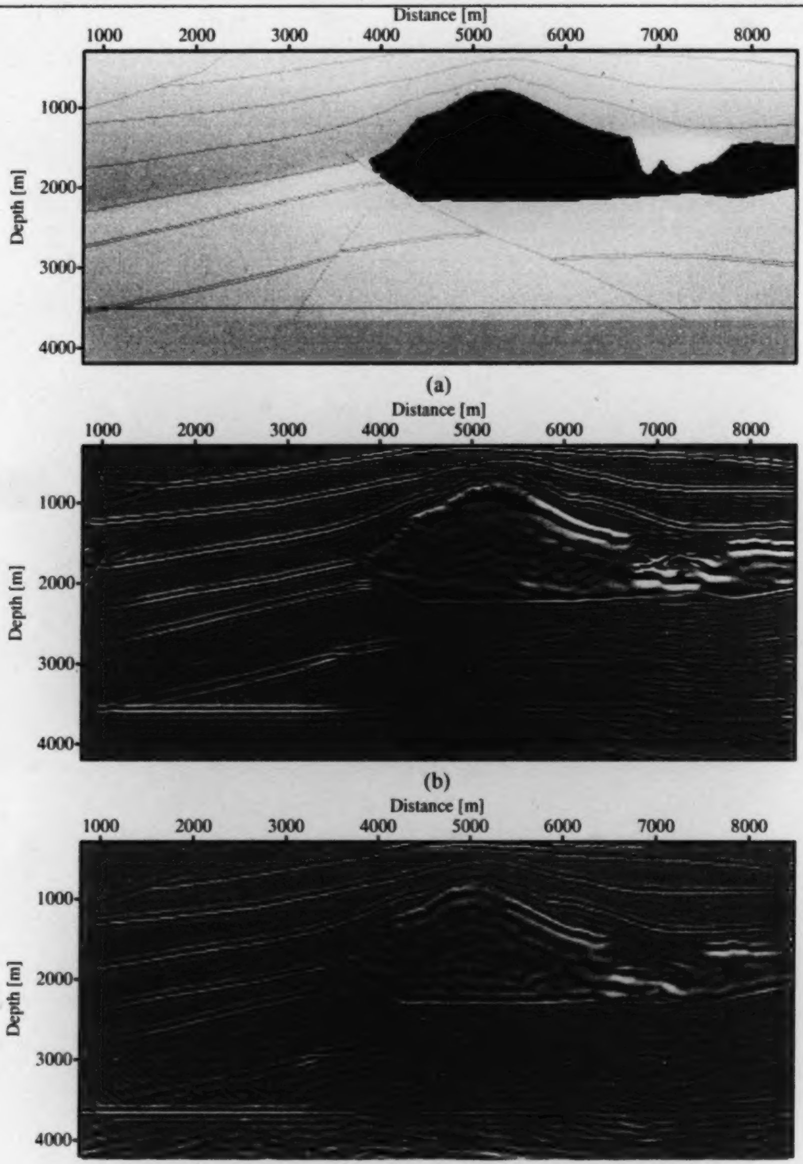


Figure 4.9: Images at $x = 6760$ m, obtained with 3-D common-azimuth migration and 3-D shot migration (online figure provided by Sandia National Laboratories). (a) Velocity profile (b) 3-D common-azimuth migration. (c) 3-D shot migration.

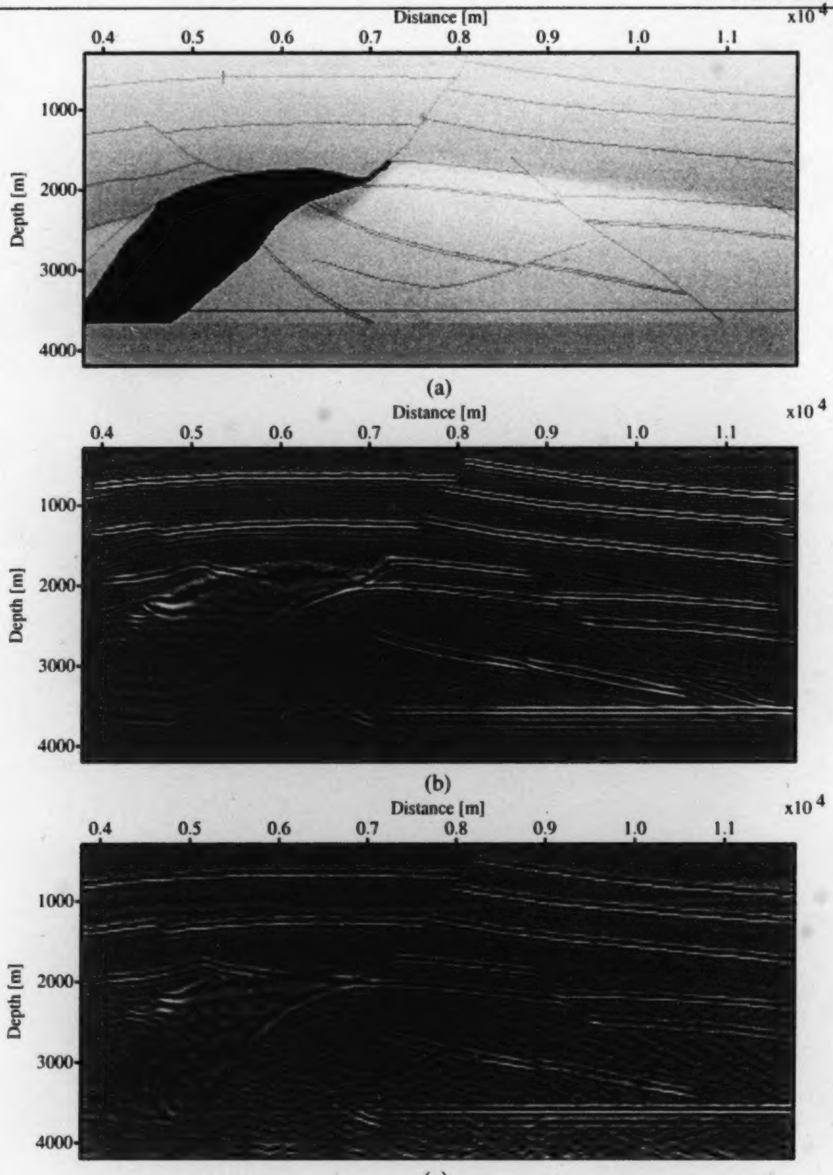


(a)

(b)

(c)

Figure 4.10: Images at $x = 8440$ m, obtained with 3-D common-azimuth migration and 3-D shot migration (online figure provided by Sandia National Laboratories). (a) Velocity profile (b) 3-D common-azimuth migration. (c) 3-D shot migration.



(a)

(b)

(c)

Figure 4.11: Images at $y = 3380$ m, obtained with 3-D common-azimuth migration and 3-D shot migration (online figure provided by Sandia National Laboratories). (a) Velocity profile (b) 3-D common-azimuth migration. (c) 3-D shot migration.

The Marmousi data set

The 2-D Marmousi data are also frequently used to test prestack migration algorithms. The data were originally produced to test velocity analysis methods. The prototype of this model was designed to simulate the geological profile through the North Quenguela Trough in the Cuanza Basin in Angola (Versteeg, 1994). The model consists of a set of slightly folded marls and carbonates. A growth fault system developed within the marls and carbonates. The velocity and density profile (see Figure 4.12) portrays the complex geological settings. As these profiles show, the model change rapidly in x and z directions, which poses a challenge to seismic imaging and velocity analysis. Especially, a good imaging algorithm should illuminate the carbonate trap at around $z = 2.5$ km and the underneath layered anticline.

The data were synthesized by an acoustic finite difference modelling method with second order accuracy. The source wavelet is a filtered real field wavelet. With blocky structure of velocity and density, the acquired data are suitable for tests of both structural imaging and AVA studies. The data were acquired with a marine data geometry by using a single cable. A total of 240 shots with 25 m spacing were simulated, each of which has 96 receivers, placed every 25 m. The offset varies from 200 m to 2575 m. The recording time is 2.896 s, and the data were resampled every 0.004 s.

The lateral velocity variation of the Marmousi model is not very large. Therefore, usually the split-step correction is good enough to preserve the structural information. As shown in Figure 4.13a, the split-step DSR migration clearly images the faults and the anticline structure. However, some events are smeared due to limited accuracy of the split-step correction. On the other hand, the split-step PSPI migration (see Figure 4.13b) helps to focus these locations. The image is cleaner than that of standard split-step DSR migration. The resolution enhancement is also evident in a comparison of the common image gathers. It is clear that the split-step PSPI migration better aligns the seismic events and helps to suppress imaging artifacts at high angles (see Figure 4.14). Note that the split-step DSR migration over corrects some events at high angles and the image gather displays a 'smile' shape. The split-step PSPI migration flattens such depth residual with more accurate velocity corrections.

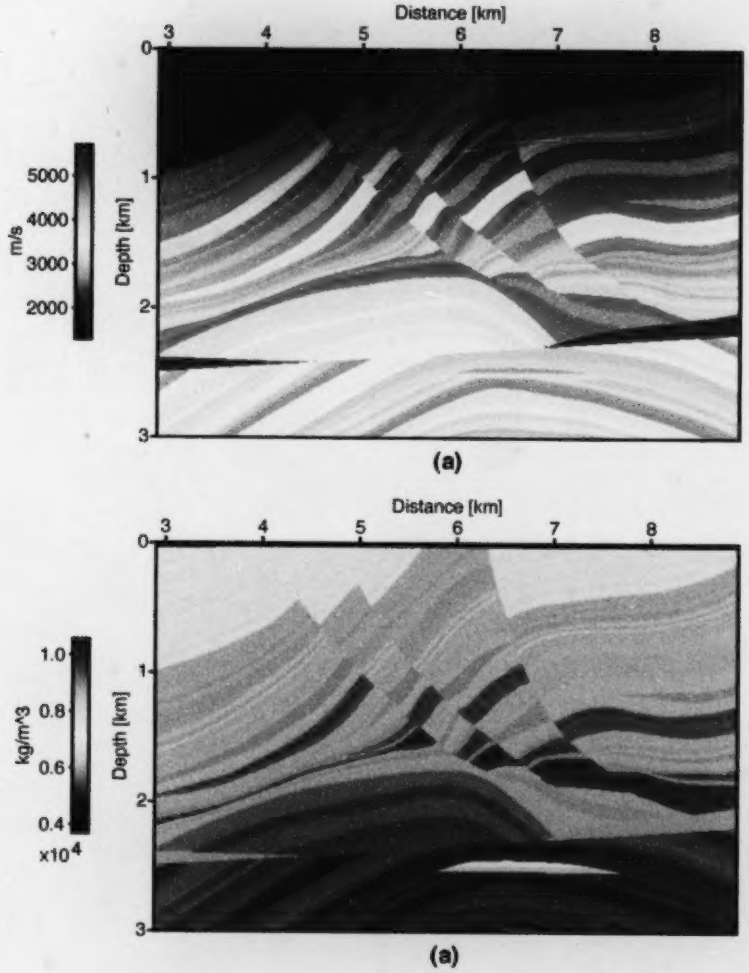
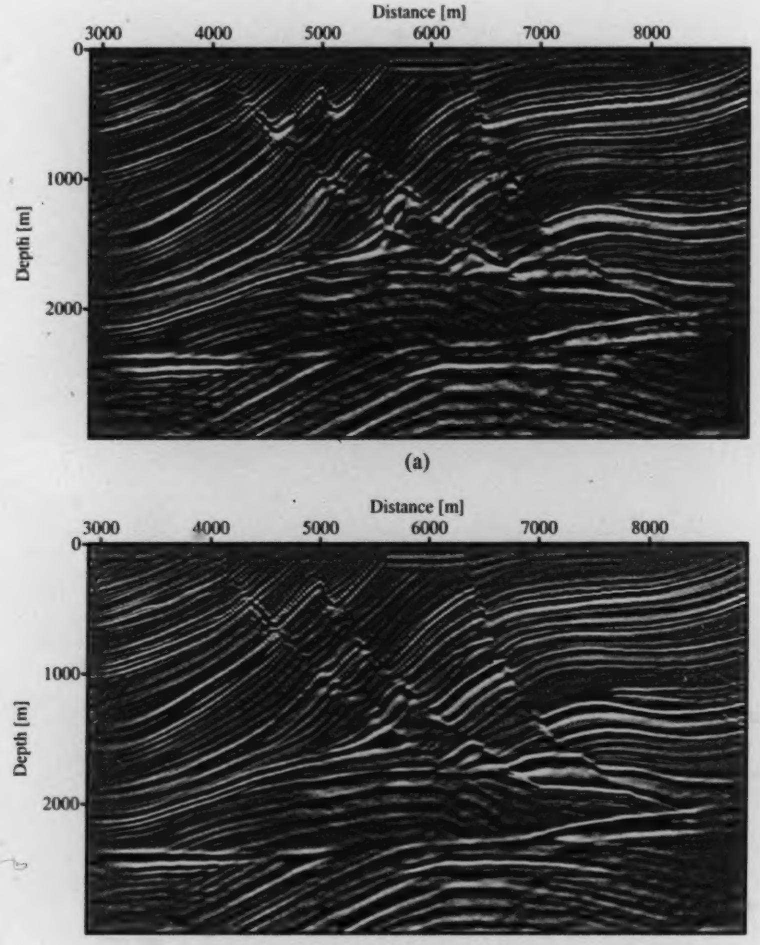


Figure 4.12: The velocity and density profiles of the Marmousi model. (a) Velocity. (b) Density. The hydrocarbon trap locates at $z = 2.5$ km and, on the top of the anticline and covered by the above folded carbonates.



(a)

(b)

Figure 4.13: Migrated images of the Marmousi data. (a) Split-step DSR migration. (b) Split-step PSPI migration. Averagely, four reference velocities were used for wavefield interpolation at each depth.

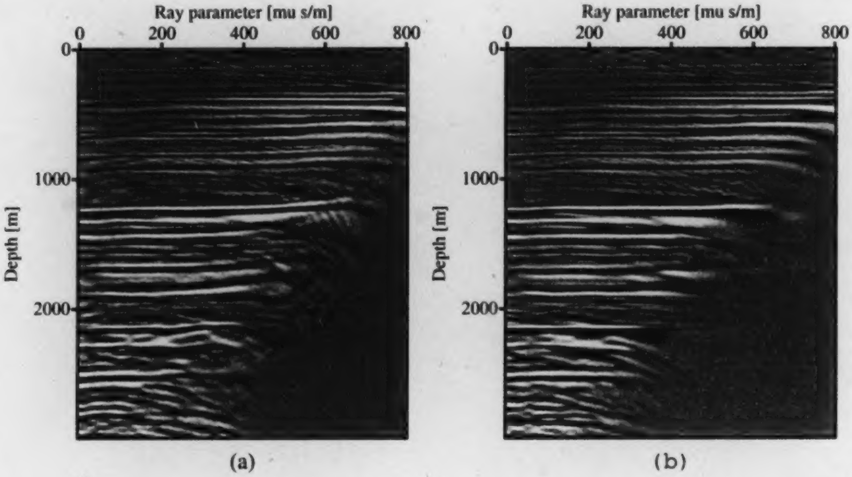

(a) (b)

Figure 4.14: Common image gathers of the Marmousi data at $x = 7500$ m . (a) Split-step DSR migration. (b) Split-step PSPI migration.

4.2 Least-squares migration

4.2.1 Least-square migration with smooth regularization

To glean the benefits of least-squares migration, I created a simple 2-D acoustic data set using a ray tracer that accounts for the correct reflector AVA, cylindrical divergence (line sources), and interface transmission losses in a laterally invariant earth model. The data were then copied to a number of inlines to simulate 3-D common-azimuth data generated by line sources. The model consists of four flat layers and a half-space. Table 4.1 describes the model in terms of velocity and density. Note that cylindrical divergence is in agreement with the assumptions made for the common-azimuth migration operator (Biondi and Palacharla, 1996). When dealing with real data, I approximately transform point sources to line sources by multiplying the data with the square root of the two-way traveltime.

Each CMP gather has 61 offsets with a spacing of 25 m. The data set consists of 10 inlines, each of which has 10 CMP gathers. Both inline and crossline CMP spacings are

Velocity (m/s)	Density (g/cm^3)	Thickness (m)
2000	2.25	500
2350	1.6	300
1900	2.3	300
2500	1.7	300
2500	2.0	Half-space

Table 4.1: Model parameters for the 3-D synthetic data.

25 m.

Noise-free incomplete data

To see if the algorithm can deal with incomplete data, I randomly removed 70% of the traces to simulate a very sparse 3-D survey.

Figure 4.15 illustrates the capability of RLSM to reconstruct 3-D seismic data. Figure 4.15a displays four CMP gathers of the original data, and Figure 4.15b shows the reconstructed CMP gathers after 15 CG iterations. The residuals (Figure 4.15c) are insignificant. This is a reassuring, albeit expected, result. Obviously, successful data reconstruction prior to conventional migration would be a more efficient alternative to least-squares migration in this case. Many schemes for wavefield reconstruction have been proposed and successfully applied. For example, Liu et al. (2003) have devised a Fourier reconstruction method that can deal with sparse data similar to the example shown in this section. However, the purpose of this study is to glean the benefits of least-squares migration without resorting to other processing techniques. This is not only of academic interest, since least-squares migration promises to achieve benefits beyond data reconstruction. Hu et al. (2001) have demonstrated that migration deconvolution, implicitly accounted for in least-squares migration, can help to sharpen the seismic image by deconvolving the migration point-spread function. The point-spread function results from the bandlimited nature of the imaging process and tends to blur the seismic image. In the real data example in the next chapter, least-squares migration does generate a higher resolution image than the conventional migration, which is in agreement with the assertion made by Hu et al. (2001).

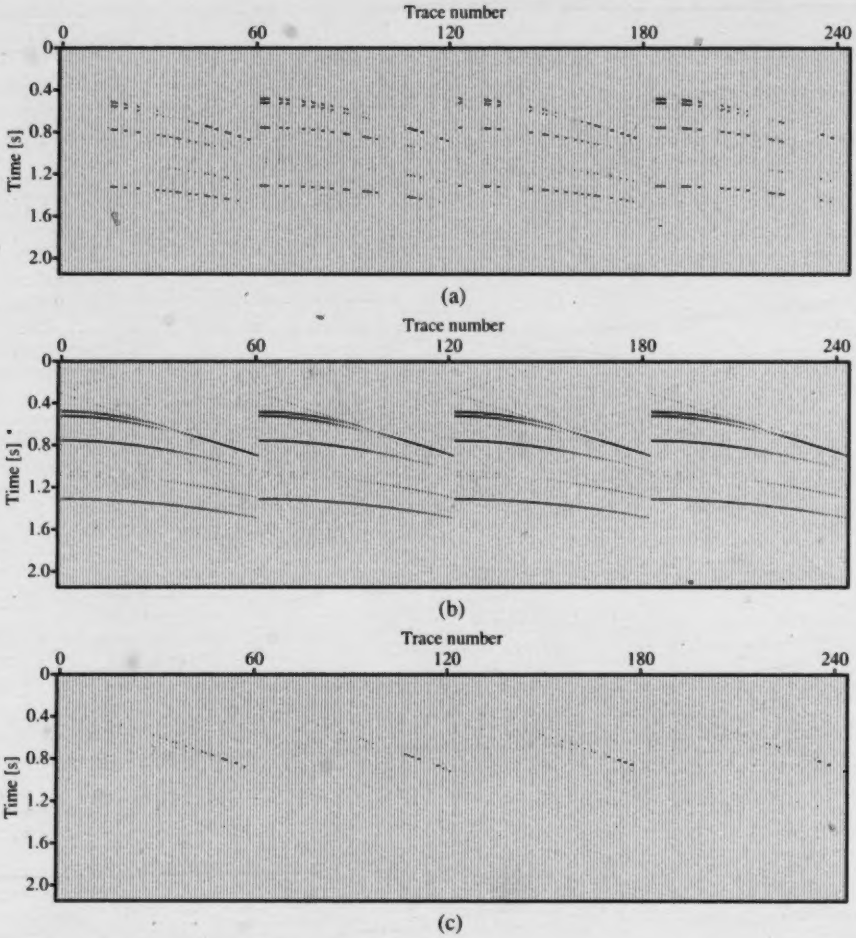


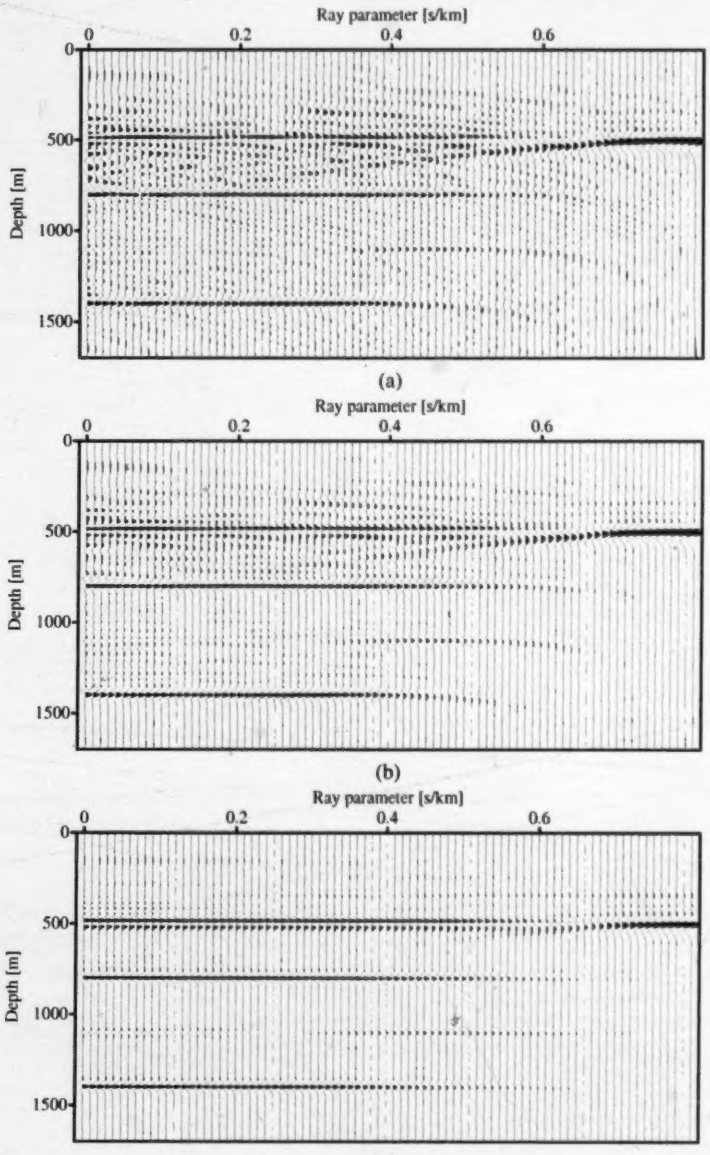
Figure 4.15: Comparison between observed and reconstructed data. (a) Four neighboring CMP gathers. Each gather has 61 offsets with 25 m spacing. (b) Reconstructed CMP gathers after 15 CG iterations. (c) Difference between the original live traces and the corresponding reconstructions.

Figure 4.16 compares common image gathers (at the same spatial location) computed via migration, RLSM after 4 CG iterations, and RLSM after 15 iterations. With more CG iterations the coherency of the CIG is gradually improved, and aliasing artifacts are further mitigated. Notice the limited aperture effect in the migration result manifesting itself as spurious tails. Part of the problem lies in the fact that the highest invertible, full bandwidth ray-parameter is determined by the upper limit of the offset wavenumber. Going beyond this limit in the radial-trace transform reduces the wavelet bandwidth. This is particularly obvious for the first event in our example. One has to be aware of this, and muting the CIGs at the upper ray-parameter limit may be advisable. However, a suppression of this effect would be preferable. For a linear approach, refer to Sava and Fomel (2003) for a more detailed discussion on angle and ray-parameter domain imaging. In this thesis, I attack this problem by introducing a sparse regularization in the depth direction in conjunct with the smooth regularization in the ray parameter direction. In a test of a wedge model further bellow, accurate amplitude can be acquired for very high angles.

Figure 4.17 compares the extracted AVA curves of the four events by two different methods. The top and bottom panels show the migration (with the imaging Jacobian correction applied) and the RLSM result, respectively. For better comparison, the smoothed migrated AVA curves are also shown in the top row. The migrated amplitudes are obviously distorted by missing data. On the other hand, RLSM retrieves the AVA closely within the invertible angle range.Notice that the effect of the imaging Jacobian is inherent in the inversion.

The data were also processed with PLSM to compare its performance with RLSM. I found that PLSM is more efficient considering the computational cost. Figure 4.18 compares calculated common image gathers by the conventional migration, RLSM (after 15 CG iterations) and PLSM (only after 6 CG iterations). It is clear that both inversion methods provide more coherent results than the conventional migration. Since the time spent on each iteration of RLSM and PLSM is almost the same (actually PLSM is slightly faster), the total amount of computational time is saved by using PLSM. As shown in Figure 4.18, the quality of the solution given by PLSM is comparable as that by RLSM. To further compare these two methods, I plot the AVA curves for all four events (see Figure 4.19). Basically, both methods provide promising amplitude within the invertible angle range.

The data fitting of RLSM and PLSM is also compared (see Figure 4.20). It can be seen that PLSM converges faster than RLSM. The algorithm starts to converge after 6 iterations.



(a)

(b)

(c)

Figure 4.16: CIGs with varying CG iterations. (a) CIG after migration. (b) CIG after 4 CG iterations. (c) CIG after 15 CG iterations.

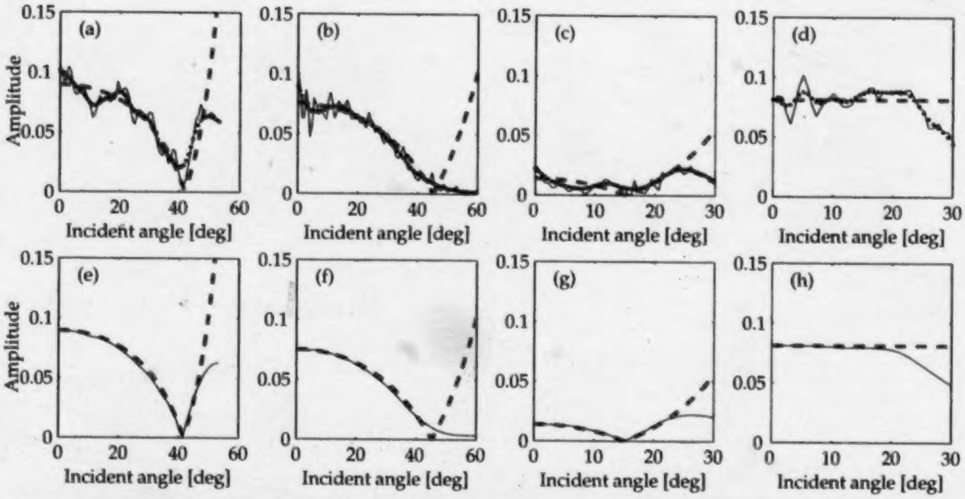
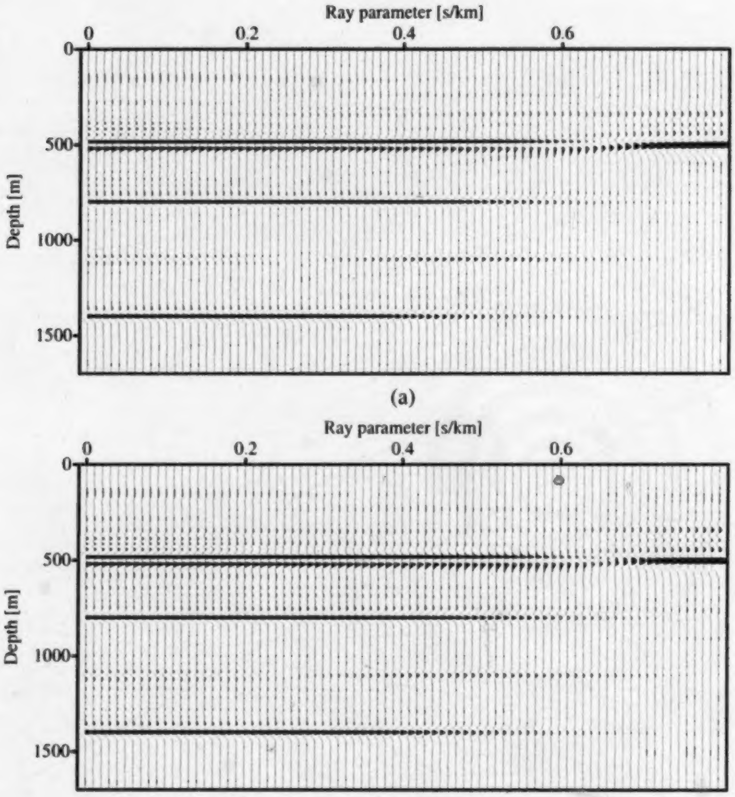


Figure 4.17: Extracted AVA curves. (a) - (d) are produced by migration. (a), (b), (c) and (d) are AVA curves for the events at depths 500 m, 800 m,1100 m and 1400 m, respectively. The red dashed lines are the theoretical AVA, the solid lines are the migration result, and the doted lines are the smoothed migration result (Hamming filtered). (e) - (h) are produced by RLSM (15 iterations). The solid lines are the RLSM inverted AVA.



(a)

(b)

Figure 4.18: CIGs by Regularized Least-Squares Migration (RLSM) and Preconditioned Least-Squares Migration (PLSM). (a) RLSM (15 iterations). (b) PLSM (6 iterations)

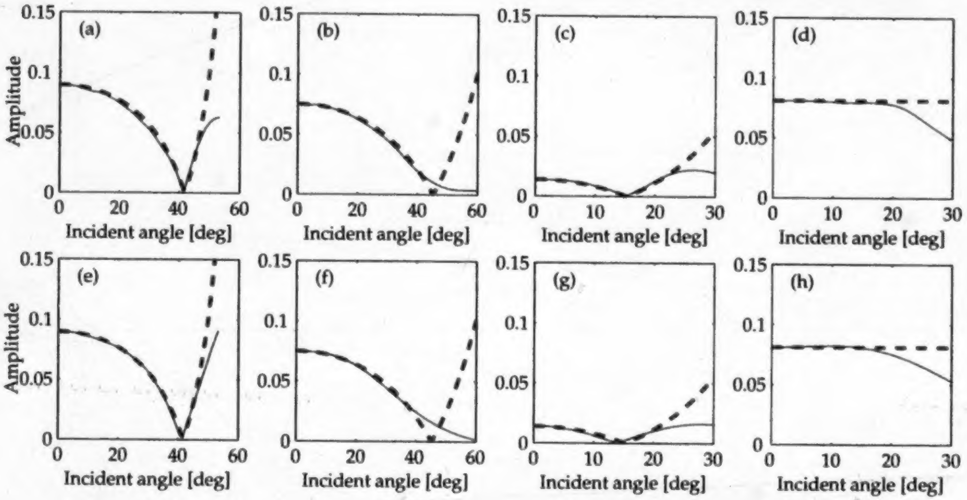


Figure 4.19: Extracted AVA curves. (a) - (d) are produced by RLSM (15 iterations). (a), (b), (c) and (d) are AVA curves for the events at depths 500 m, 800 m,1100 m and 1400 m, respectively. The red dashed lines are the theoretical AVA, the blue solid lines are the migration result. (e) - (h) are produced by PLSM (6 iterations). The red dashed lines are the theoretical AVA, and the blue solid lines are the PLSM inverted AVA.

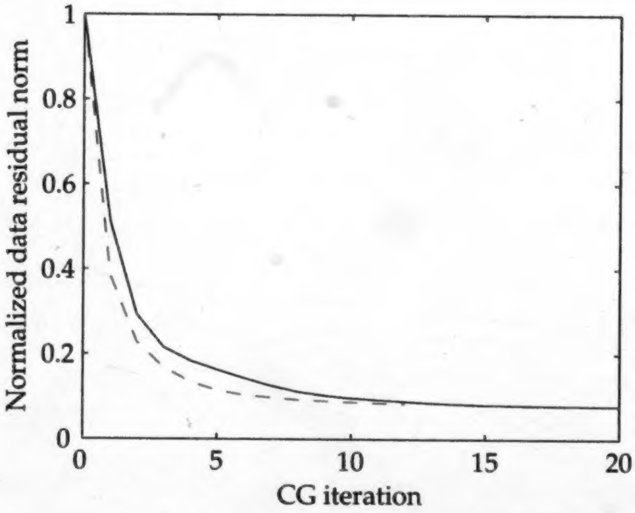


Figure 4.20: Data misfit of two inversion methods. Blue solid: RLSM. Red dashed: PLSM.

Noisy incomplete data

As shown in the previous chapter, regularized least-square inversion is efficient in suppressing noise in the deconvolution problem. In the example of inverting the smooth model, the smooth regularization was used for one single trace. The point was to conduct a fast test of the inversion prototype. This is not realistic for a single seismic trace. However, the idea of forcing smoothness can be adopted for prestack angle domain imaging problems. The method can be understood as a multi-channel inversion, in which neighboring coherent useful information are enforced. To verify this point, I added some white noise with moderate signal-to-noise ratio ($S/N = 10$) to the previously synthesized common azimuth data and re-ran the program. For efficiency, I only tested the PLSM algorithm. It has been shown that (Kuehl, 2002) RLSM is not sensitive to this level of white noise for moderate signal-to-noise ratios. Please note that I assume that coherent noise like multiples have been properly removed before applying least-squares migration. Therefore, the influence of coherent noise is not examined here.

Figure 4.21 compares the common image gathers produced by the conventional migration and PLSM. Obviously, the result of the conventional method is not easy to use due to the contained severe aliasing and noise. It is difficult to identify the coherent pattern and pick the events. Especially the third event with small amplitude is immersed in the noise. The picked AVA (see Figure 4.22a-d)deviates significantly from the theoretical value, and simple smoothing of the amplitude does not help much. On the other hand, PLSM efficiently suppresses the white noise and aliasing artifacts caused by data missing. All four events are resolved in a crisp manner. The amplitude is still accurate as shown in Figure4.22e-h though the fitting angle range shrinks slightly.

The evolution of the data fitting in the PLSM method is shown in Figure 4.23. It can be seen that the data fitting improves during the first four CG iterations, then deteriorates. The reason is two-fold. First, the preconditioning without trade-off control entails dealing with a rank-deficient problem. Full convergence is not advised if one seeks a good solution. Second, with too many iterations, the algorithm tries to fit the noise, and therefore, the creation of artifacts representing date noise mapped to the model space and not real desirable features of our inverted model. Therefore, an optimal solution should be selected from the first few iterations. After examining the data misfit curve, I found that the solution at the fourth iteration best honors the observation.

The efficiency of PLSM is further confirmed in a comparison of the original data and the reconstructed data (see Figure 4.24). The reconstructed data are very clean and the coherent information is preserved nicely. The residual, or the predicted noise, consists mainly the random noise. The difference between this predicted and the true noise is slight. As portrayed in Figure 4.25, no significant valuable information is contained in the predicted noise panel. Only far offset energy in the the first event leaked in the residual panel.

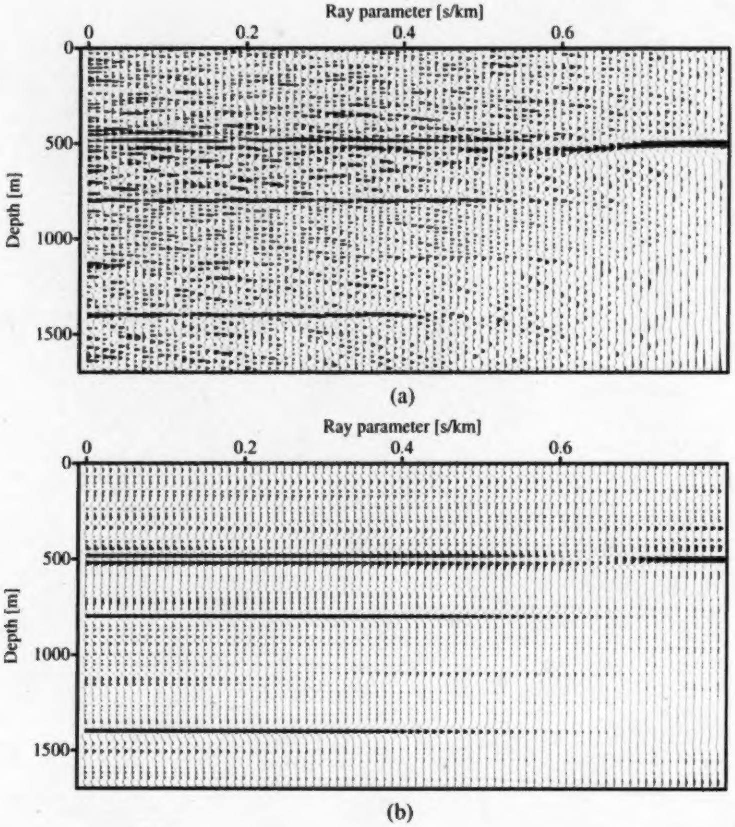


Figure 4.21: CIGs by migration and Preconditioned Least-Squares Migration (PLSM). (a) Migration. (b) PLSM (4 iterations). The input data were randomly decimated (70% traces removed) and contaminated by some significant white noise ($S/N = 10$).

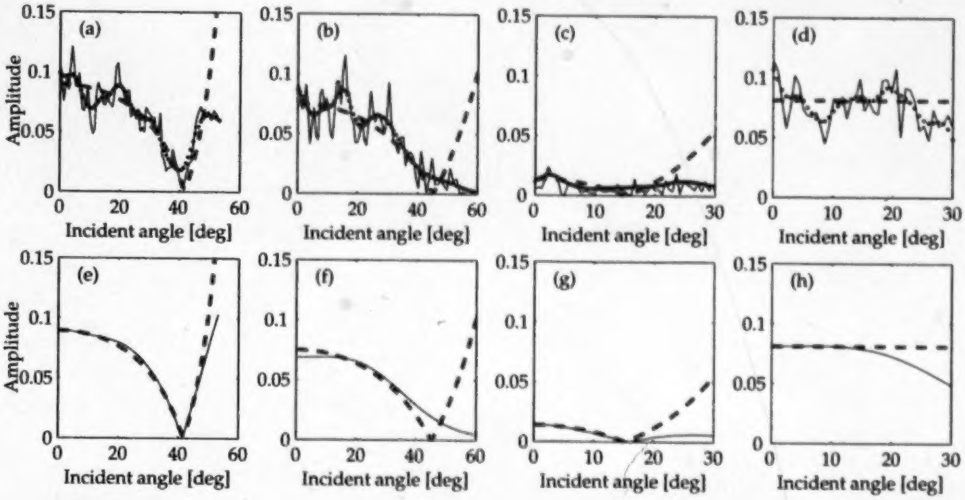


Figure 4.22: Extracted AVA curves of the noisy incomplete data. (a) - (d) are produced by migration. (a), (b), (c) and (d) are AVA curves for the events at depths 500 m, 800 m, 1100 m and 1400 m, respectively. The red dashed lines are the theoretical AVA, the solid lines are the migration result, and the doted lines are the smoothed migration result (Hamming filtered). (e) - (h) are produced by PLSM (4 iterations). The solid blue lines are the PLSM inverted AVA.

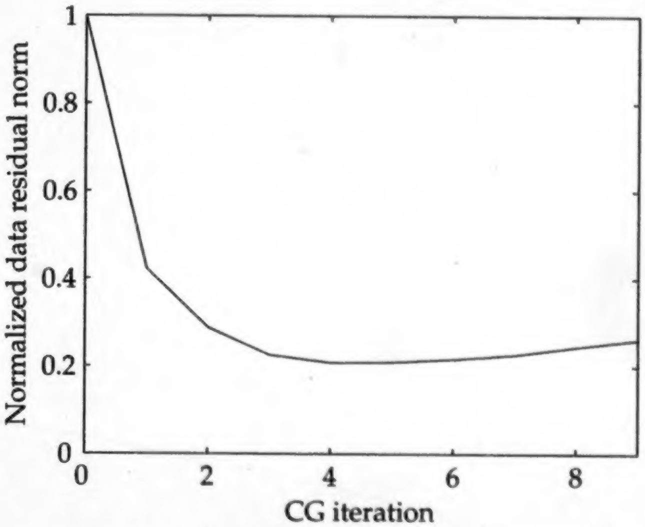


Figure 4.23: Data misfit of the preconditioned least-squares migration applied to the noisy incomplete data. The optimal solution is chosen at the fourth iteration.

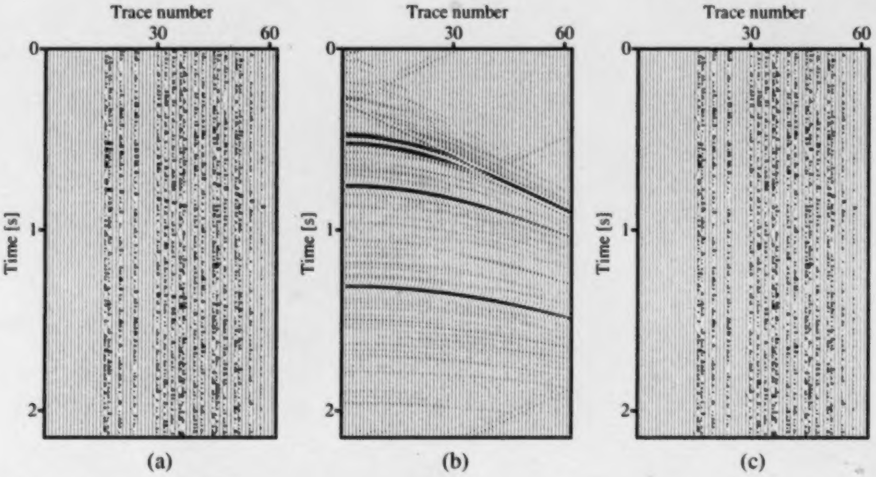


Figure 4.24: Comparison of the observed, reconstructed data and the residual (a) The original data. (b) The reconstructed complete data. (c) The difference between the live traces and the corresponding reconstructed traces

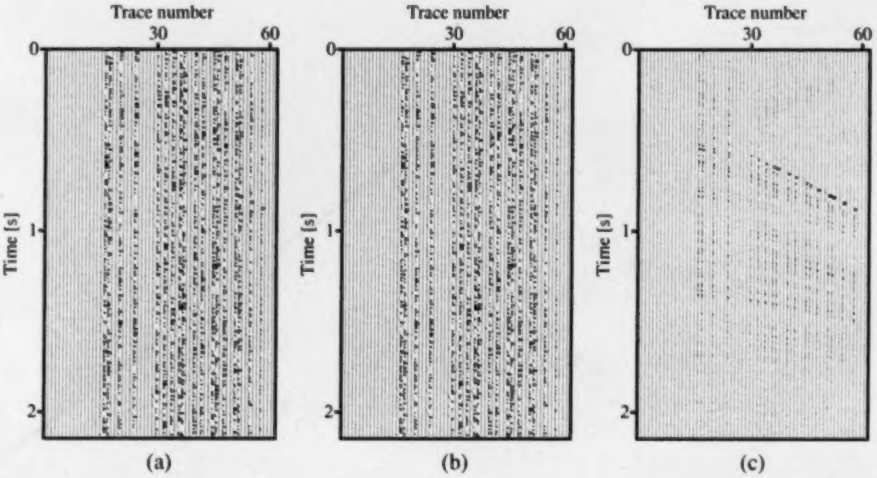


Figure 4.25: Comparison of the predicted noise, the real additive noise and the residual. (a) The predicted noise. (b) The real additive noise. (c) The residual.

4.2.2 Sparse least-squares migration

As shown in the *toy* example of multi-channel deconvolution in Chapter 3, combining sparse and smooth regularization can lead to high resolution in two directions. The goal is to acquire sparseness in the depth direction and smoothness in the ray parameter direction. Now I test the idea with a 2-D wedge model. A 3-D test on the field data is shown in the next chapter. The 2-D wedge model gives an economical way to study the behavior of this method. With tapering thickness between two neighboring reflectors, the model is useful for examining the influence of inversion on the tuning effect, which is rooted in the band-limited nature of seismic data. Figure 4.26 portrays the geometry of the wedge model. The thickness changes gradually from 0 m to 100 m with a step of 5 m. The midpoint spacing is 25 m. The step in the vertical direction is 5 m.

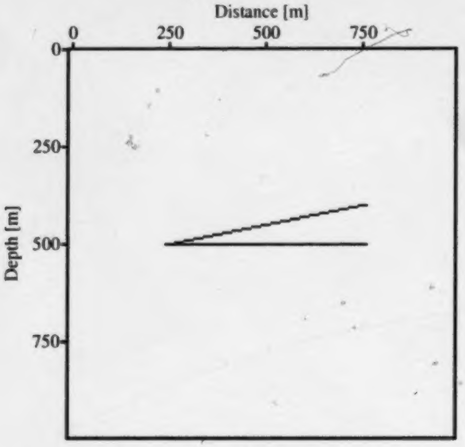


Figure 4.26: The geological structure of the wedge model.

Two sets of data were prepared to examine the effects of wavelet mismatch. The first dataset was prepared using the following forward modeling operator:

$$\mathbf{L}: \qquad \mathbf{d} = \mathbf{L}_\omega \mathbf{A}\mathbf{m}, \tag{4.1}$$

where $\mathbf{d}$ is the seismic data, $\mathbf{L}_\omega$ is a band-limited wave propagator, $\mathbf{A}$ is the adjoint of the radial trace transform, and $\mathbf{m}$ is the AVP image gather. The AVP of the first of the

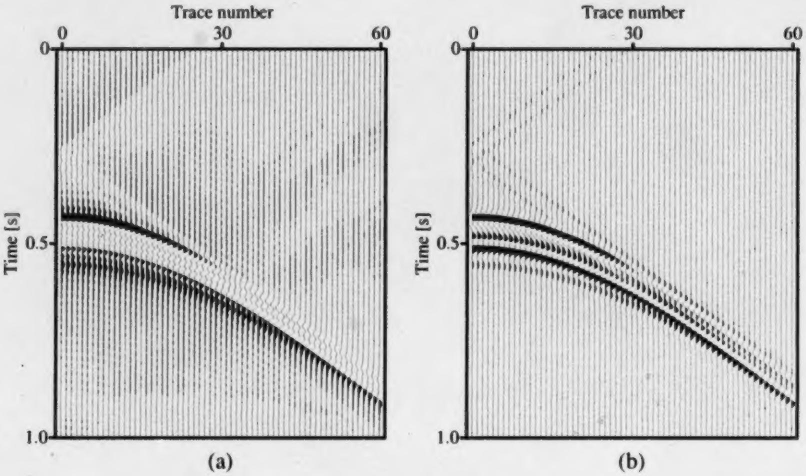


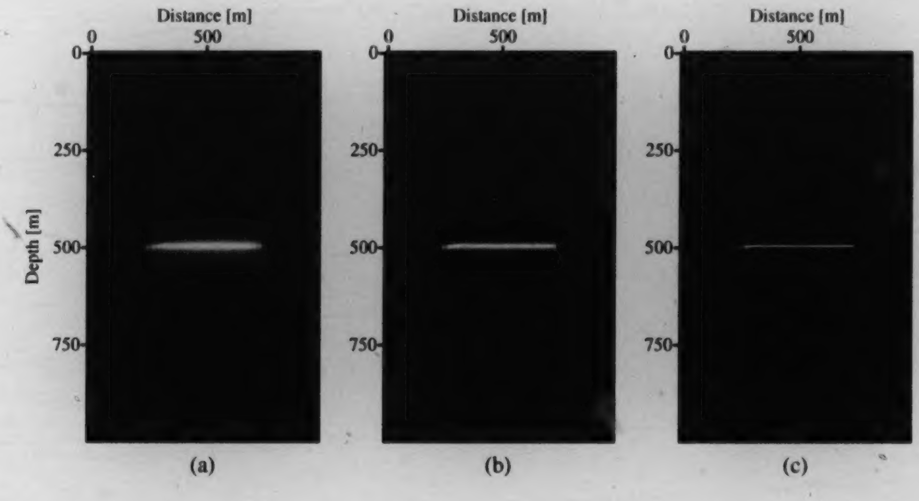
Figure 4.27: Synthesized common midpoint gathers at midpoint $x = 500$ m of the wedge model. (a) Data without the source signature. (b) Data with the source signature.

previous model is used for the tilted reflector of the wedge model, The AVP of the flat reflector at depth $z = 500$ m is set to a negative constant. Compared with equation 2.171, the above equation ignores the source signature $\mathbf{W}_s$. The second dataset was prepared considering the source signature, exactly following equation 2.171. A Ricker wavelet is used to represent the source signature. Both datasets were produced for a frequency range, from 0 to 50 Hz, and sampled every 0.004 second in the time domain. The recording time is 2.148 s (538 samples each trace). The CMP gathers at midpoint $x = 500$ m are displayed in Figure 4.27. The tuning effects are evident in both datasets. Hereafter, the dataset without source signature is denoted as dataset 1, and the one with source signature is denoted as dataset 2.

Noise free dataset 1

I processed dataset 1 by three methods: conventional migration (the adjoint of the modeling operator), preconditioned least-squares migration (PLSM) and the sparse least-squares migration (SLSM).

Figure 4.28 shows the stacked images. The result of the adjoint is quite blurry since



(a) (b) (c)

Figure 4.28: Stacked images of the wedge model, inverted from noise-free dataset 1. (a) Migration. (b) PLSM. (c) SLSM.

the algorithm is not capable of reconstructing high frequencies absent from the data. On the other hand, both PLSM and SLSM are efficient at recovering the structural images. Especially, the SLSM algorithm has produced a highly resolved image. This is a consequence of using a sparseness constraint that attempts to collapse the band-limiting seismic wavelet into a broad-band impulsive signal.

Figure 4.29a-c displays a zoomed view of three common image gathers produced by these methods. The SLSM method has the ability of suppressing the sidelobes introduced by the band-limited nature of the data. To complete the analysis, I have extracted the amplitude of the tilted event and plotted AVA curves for the three methods in Figure 4.29d. I can observe that both PLSM and SLSM are able to preserve the amplitude response of the reflection. SLSM provides a larger fitting angle range than PLSM.

Figure 4.30 compares the data misfit of two inversion methods. The PLSM starts to converge at the seventh iteration. As shown in Figure 4.29d, the method provides accurate amplitude for an angle range between 0 and 40 degrees. However, the structural image is not satisfactory due to the tuning effects. The SLSM method needs four itera-

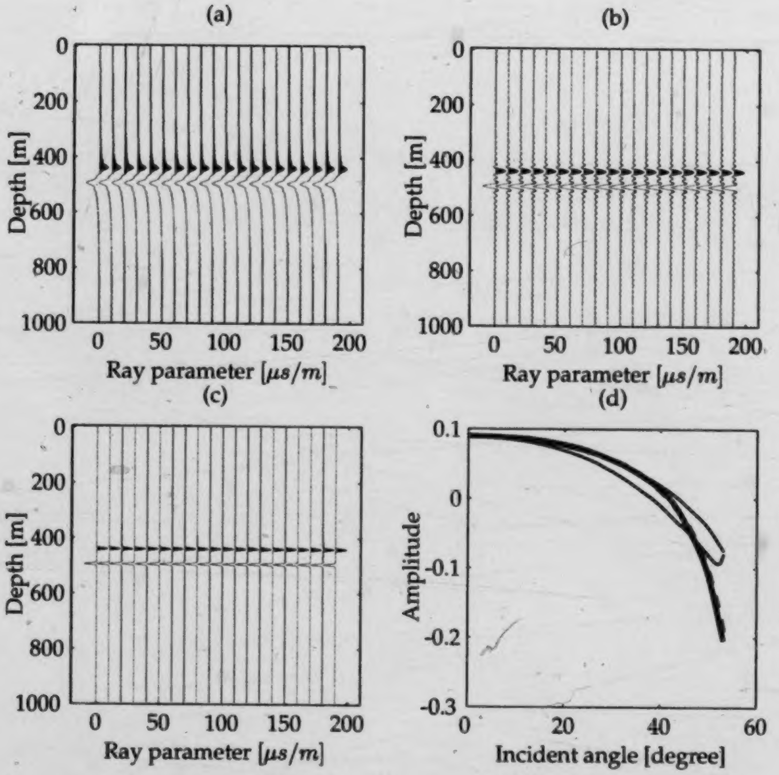


Figure 4.29: Common image gathers (CIGs) and AVA curves at $x = 500$ m for the wedge model. (a) Migration. (b) Preconditioned least-squares migration (PLSM). (c) Sparse least-squares migration (SLSM). (d) AVA curves for the first event. Red dashed: the theoretical curve. Green: migration. Blue: PLSM. Black: SLSM.

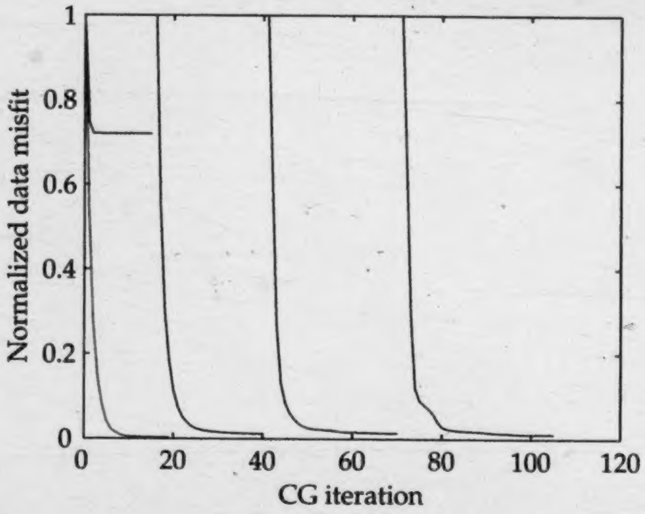
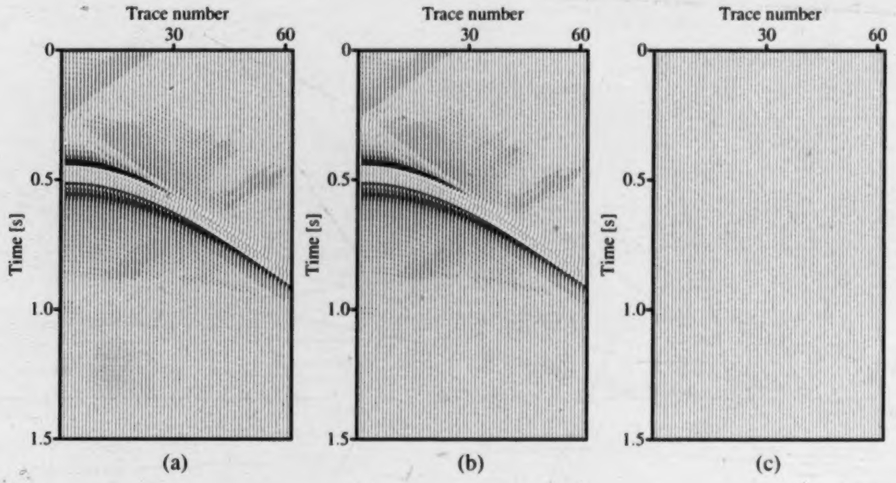


Figure 4.30: Data misfit of two inversion methods. Blue: PLSM. Red: SLSM.

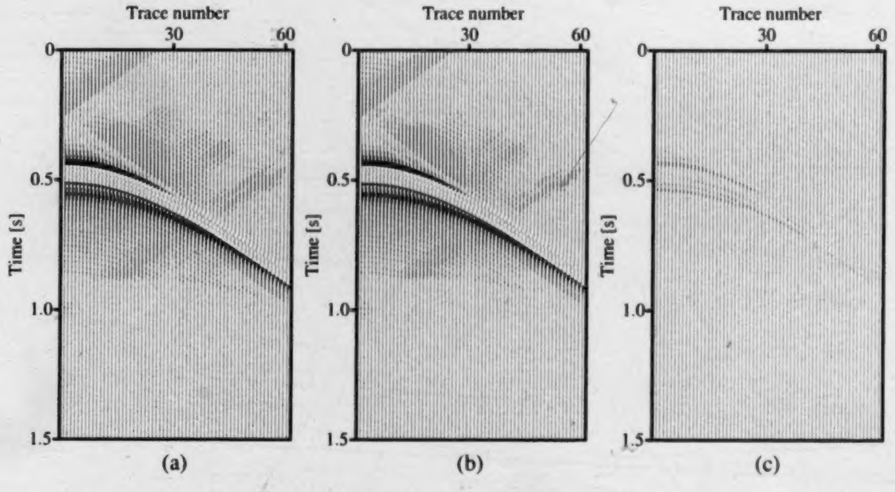
tions of IRLS algorithm to acquire convergent sparse solution. The data fitting improves with more IRLS iterations. In this test, the amplitude completely fits the theoretical AVA value. The wavelet has collapsed completely with the sparse regularization.

Figure 4.31 compares the original data, reconstructed data by PLSM and the residual. Again, we can see the efficiency of PLSM at fitting the data. The residual is very clean. The data fitting of SLSM (see Figure 4.32) is also impressive though slightly inferior to PLSM. A small amount of data are not fitted.



(a) (b) (c)

Figure 4.31: The original data, reconstructed data by PLSM and the residual at $x = 500$ m. (a) Original. (b) Reconstructed. (c) Residual.



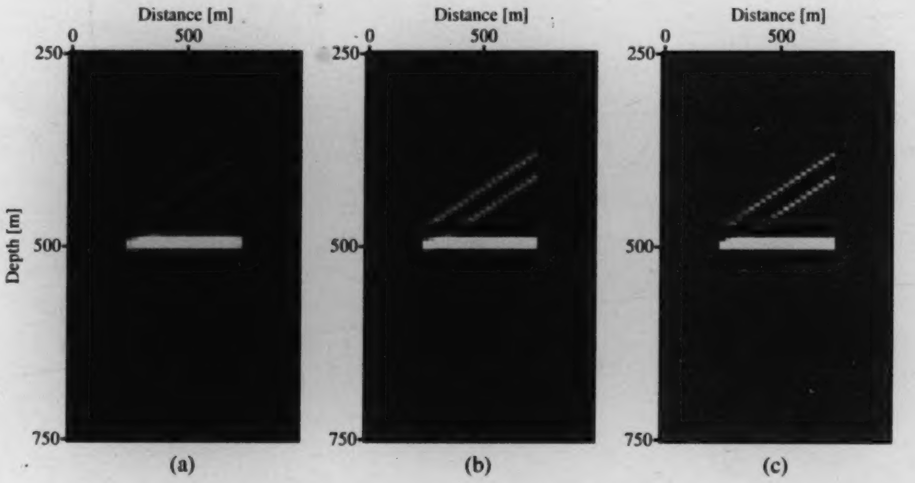
(a) (b) (c)

Figure 4.32: The original data, reconstructed data by SLSM and the residual at $x = 500$ m. (a) Original. (b) Reconstructed. (c) Residual.

Noise-free dataset 2

Similar comparison was conducted on the noise-free dataset 2. Since the source signature is included in the data, but it is not taken into account in the inversion operator, we can not expect so much high resolution as the previous experiment provides. On the other hand, if we include the source simulation in the operator, the bar can also be lifted.

The first test is to examine the performance of least-squares migration without considering the source signature. Figure 4.33 compares the stacked images by migration, PLSM and SLSM. It can be seen that both PLSM and SLSM provide higher solution than the conventional migration. SLSM better resolves the two events, but it is not so good as the previous example.The reason is that the source wavelet is not encoded in the modeling/adjoint pairs. As a result, the algorithm mainly aims at inverting the implicit wavelet of the wave propagator (L_ω)and removing small artifacts.



(a) (b) (c)

Figure 4.33: Stacked images of the wedge model, inverted from noise-free dataset 2. The source wavelet is not considered. (a) Migration. (b) PLSM (4 iterations). (c) SLSM (3 IRLS iterations). The source signature is not considered in PLSM and SLSM.

The detail comparison of the common image gathers also supports the above explanation. The PLSM does less than in the previous example to improve the vertical solution though some de-blurring effects can be observed in the stacked image. In SLSM, the wavelet is slightly suppressed. The amplitude of PLSM and SLSM is better than the conventional migration (Figure 4.34). Furthermore, SLSM gives better result at high angles than PLSM.

The second test is to encode the source signature (wavelet) in three methods: migration, PLSM and SLSM. In migration, the adjoint of the wavelet is used to correct the wavefield before downward continuation of the algorithm. This is just to give a complete comparison of the adjoint and the inversion operators. The correction does not invert the source wavelet, and the amplitude in the vertical direction is turned off as seen in Figure 4.35a. On the other hand, iteratively applying source wavelet and its adjoint in PLSM and SLSM helps to suppress the tuning effects (See Figure 4.35b and c). With better match of the operator, SLSM provides both exact structure and angle dependent amplitude (Figure 4.35c and d).

The observed resolution discrepancy of various methods can be explained by analyzing the least-squares problem. For simplicity, I first ignore the regularization and start from the linear system defined by equation 2.171 as bellow:

$$\mathbf{L}: \qquad \mathbf{d} = \mathbf{W}_s \mathbf{L}_\omega \mathbf{A}\mathbf{m}. \tag{4.2}$$

By using the pseudo inverse of the wavelet $\mathbf{W}_s^\dagger$, the modeling system can be rewritten as:

$$\begin{aligned} \mathbf{W}_s^\dagger \mathbf{d} &= \mathbf{L}_\omega \mathbf{A}\mathbf{m} \\ \tilde{\mathbf{d}} &= \tilde{\mathbf{L}}\mathbf{m}, \end{aligned} \tag{4.3}$$

where $\tilde{\mathbf{d}} = \mathbf{W}_s^\dagger \mathbf{d}$, and $\tilde{\mathbf{L}} = \mathbf{L}_\omega \mathbf{A}$. The least-squares solution of the problem is

$$\mathbf{m} = (\tilde{\mathbf{L}}'\tilde{\mathbf{L}})^{-1}\tilde{\mathbf{L}}'\tilde{\mathbf{d}}. \tag{4.4}$$

On the other hand, the solution without considering the source wavelet is

$$\tilde{\mathbf{m}} = (\tilde{\mathbf{L}}'\tilde{\mathbf{L}})^{-1}\tilde{\mathbf{L}}'\mathbf{d}. \tag{4.5}$$

Comparing above two solutions, I find the only difference exists in the last term, the data vector. The second solution is a smeared version due to the source signature. The solution can be probably improved by applying a bandlimited inverse to the observation before the inversion. In the Fourier domain, this may lead to instability. Consequently, the least-squares problem would try to fit oscillatory data, which is risky. For this concern, a better strategy is to encode the source wavelet in the inversion and stabilize the algorithm by regularizing the model, for example, force the solution to be smooth or sparse. The idea is fully employed in this thesis.

Noisy dataset 1

The SLSM algorithm is not sensitive to white noise either. I added some moderate noise ($S/N = 10$) and ran migration, PLSM and SLSM. The resulting common image gathers and AVA curves are shown in Figure 4.36. We can see that SLSM is the most robust method by properly imaging the simple structure and accurately retrieving the amplitude signature.

Both PLSM and SLSM are efficient at reconstructing the data (see Figure 4.37 and Figure 4.38). There is almost no coherent signal in the residual.

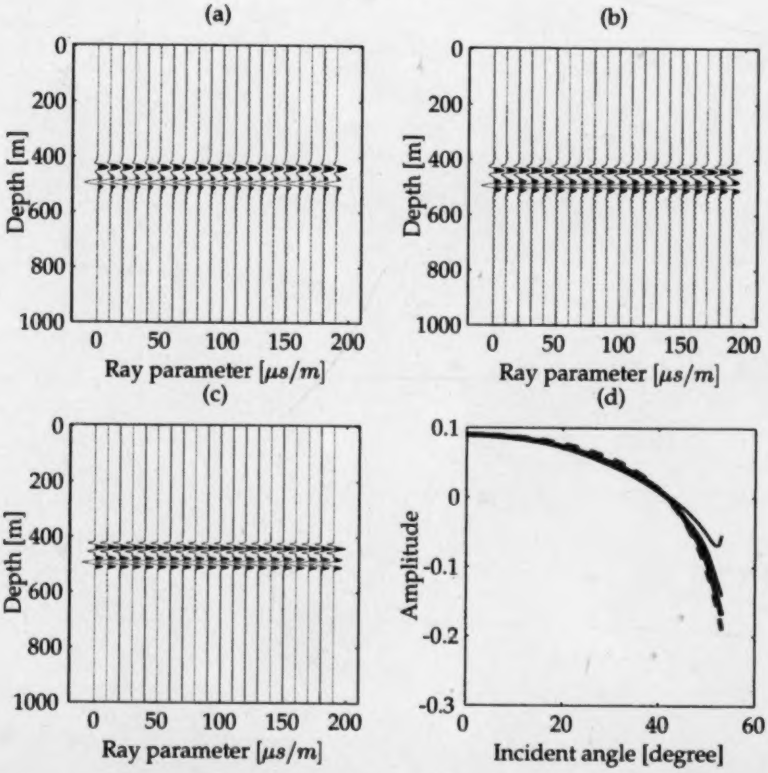


Figure 4.34: Common image gathers (CIGs) and AVA curves of the noise-free dataset 2 at $x = 500$ m of the wedge model. The source wavelet is not considered. (a) Migration. (b) Preconditioned least-squares migration (PLSM, 4 iterations). (c) Sparse least-squares migration (SLSM, 3 IRLS iterations). (d) AVA curves for the first event. Red dashed: the theoretical curve. Green: migration. Blue: PLSM. Black: SLSM.

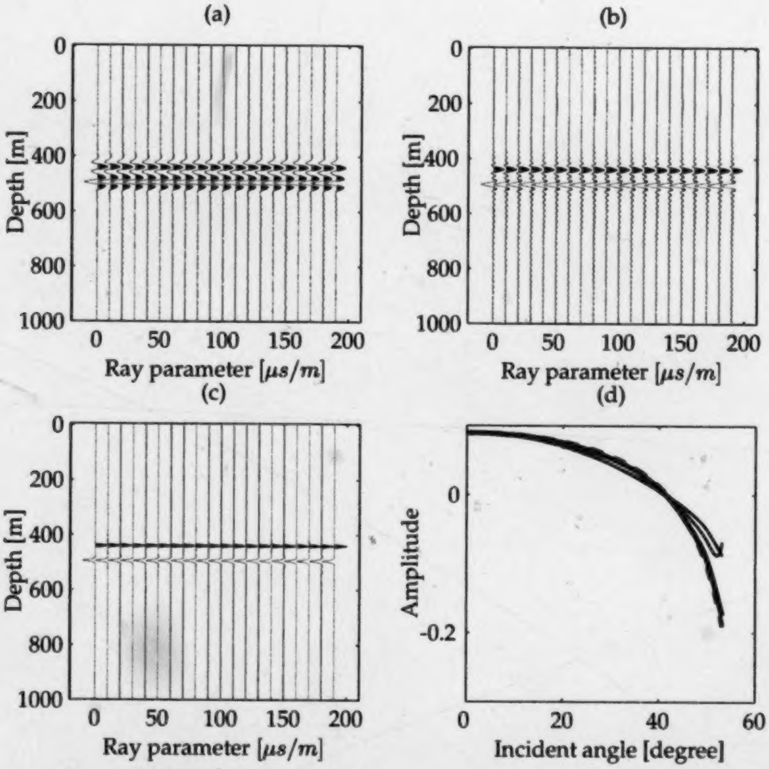


Figure 4.35: Common image gathers (CIGs) and AVA curves at $x = 500$ m of the wedge model. The source wavelet is considered (a) Migration. (b) Preconditioned least-squares migration (PLSM, 4 iterations). (c) Sparse least-squares migration (SLSM, 4 IRLS iterations). (d) AVA curves for the first event. Source signature is considered in all methods. Red dashed: the theoretical curve. Green: migration. Blue: PLSM. Black: SLSM.

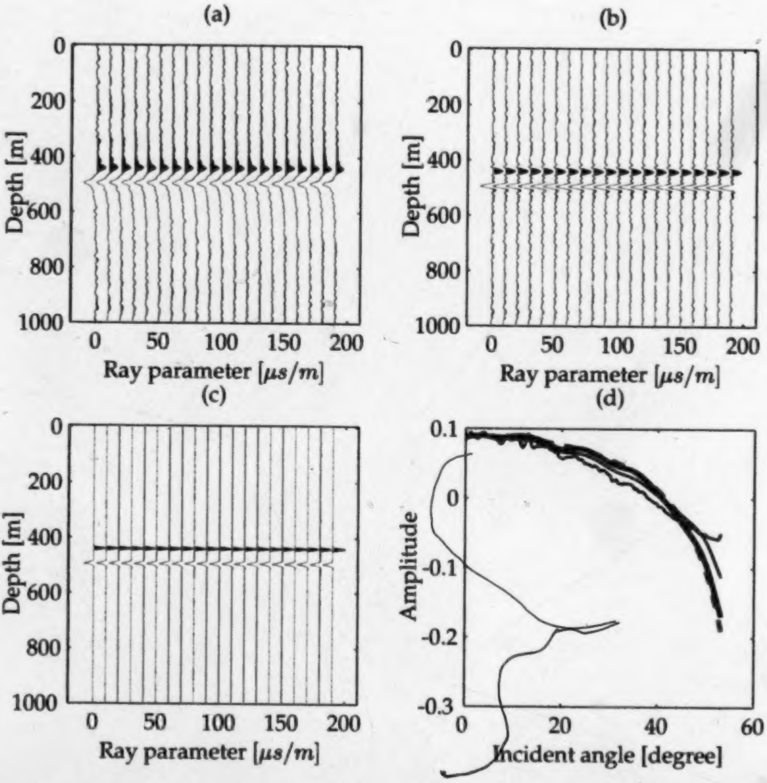
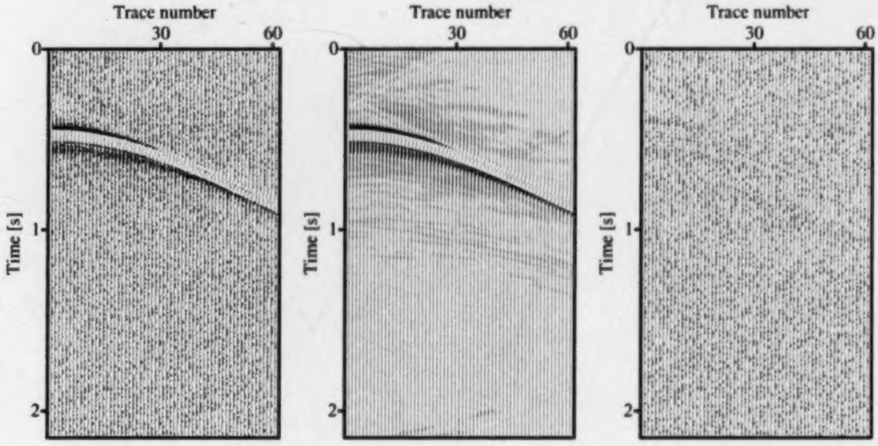
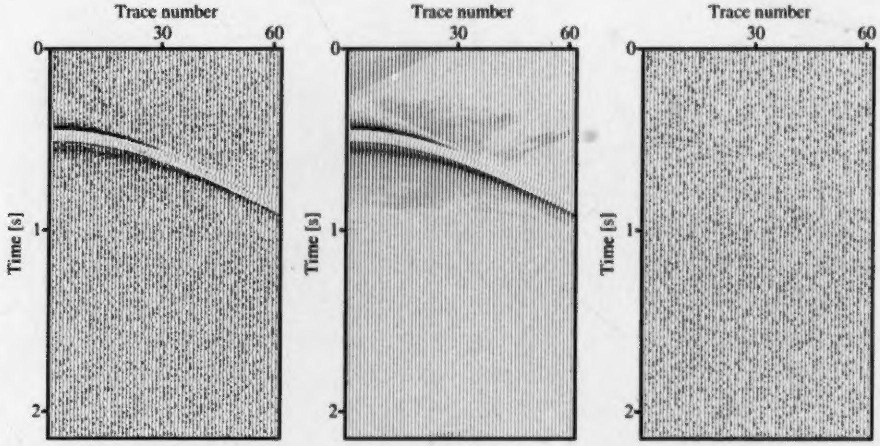


Figure 4.36: Common image gathers (CIGs) and AVA curves of the noisy dataset 1 at $x = 500$ m of the wedge model. (a) Migration. (b) Preconditioned least-squares migration (PLSM, 5 iterations). (c) Sparse least-squares migration (SLSM, 4 IRLS iterations). (d) AVA curves for the first event. Red dashed: the theoretical curve. Green: migration. Blue: PLSM. Black: SLSM.



(a) (b) (c)

Figure 4.37: Comparison between observed, reconstructed data and the residual (a) The original data. (b) The reconstructed complete data (PLSM, 6 iterations). (c) The residual (predicted noise)



(a) (b) (c)

Figure 4.38: Comparison between observed, reconstructed data and the residual. (a) The original data. (b) The reconstructed complete data (SLSM, 4 IRLS iterations). (c) The residual (predicted noise)

Marmousi model

To see if SLSM can handle complex structures and data acquisition footprint, I applied the algorithm to a re-sampled version of the Marmousi dataset. I carried out data decimation by removing 70% of traces of the Marmousi data. The decimation was done in a random manner. Migration and PLSM were also applied for comparison. The split-step PSPI operator was used in all methods to decrease operator mis-match. In migration, only the adjoint of the modelling operator was applied. In PLSM, four iterations of preconditioned CG were required to acquire reasonable data fitting, which cost about 8 times of the expense for migration. Finally, SLSM involved 3 iterations of IRLS algorithm, each of which cost 6 iterations of preconditioned CG algorithm. Therefore, the total cost of SLSM in this case is about 36 times of the conventional migration. Since I did not know the source wavelet, it is not included in the SLSM processing.

To compare the image quality, I extracted the common image gathers at $x = 7500$ m from the results given by the three methods. The wiggle plots in Figure 4.39 portray the general view of the inverted local model. It is clear that a many artifacts are present in the common image gather obtained with the migration algorithm. These artifacts are substantially removed from the images obtained with PLSM and SLSM. The latter cleans up further the image gather by suppressing sidelobes. A better comparison is given in the detail image of first few ray parameters (see Figure 4.40). For the purpose of comparison, I calculated the reflectivity series by using the true velocity and density model. A side-by-side comparison confirms that the SLSM has properly reconstructed the model. I observe again, as in the previous wedge example, an important attenuation of ringing arising from the band-limiting wavelet in the data. To evaluate the amplitude preserving properties of our algorithm, I have obtained AVA curves for the event at depth $z = 800$ m. The amplitude response obtained via the migrated image is difficult to extract due to sampling artifacts. The inverted AVA responses (PLSM and SLSM), on the other hand, are in good agreement with the theoretical value.

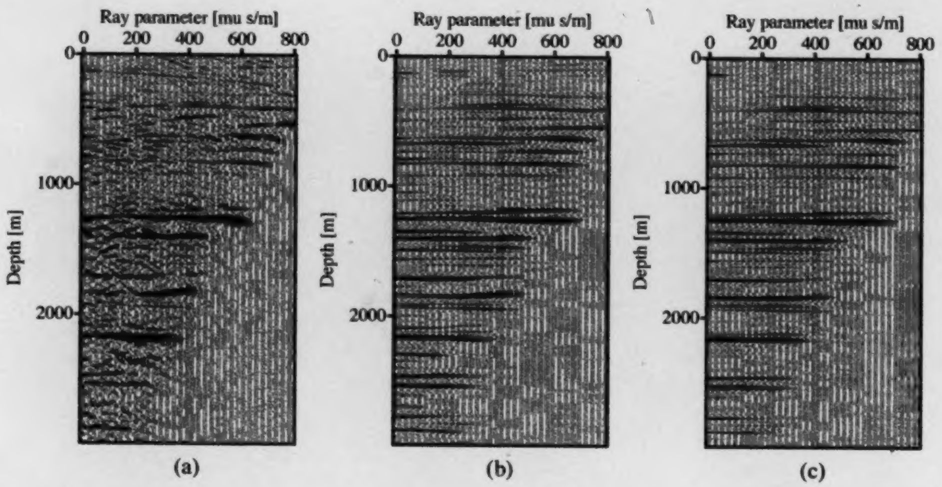


Figure 4.39: Common image gathers of migration, the Preconditioned Least-Squares Migration (PLSM) and the Sparse Least-Squares Migration (SLSM) for the incomplete Marmousi data (70% traces randomly removed), at $x = 7500$ m. (a) Migration. (b) PLSM. (c) SLSM

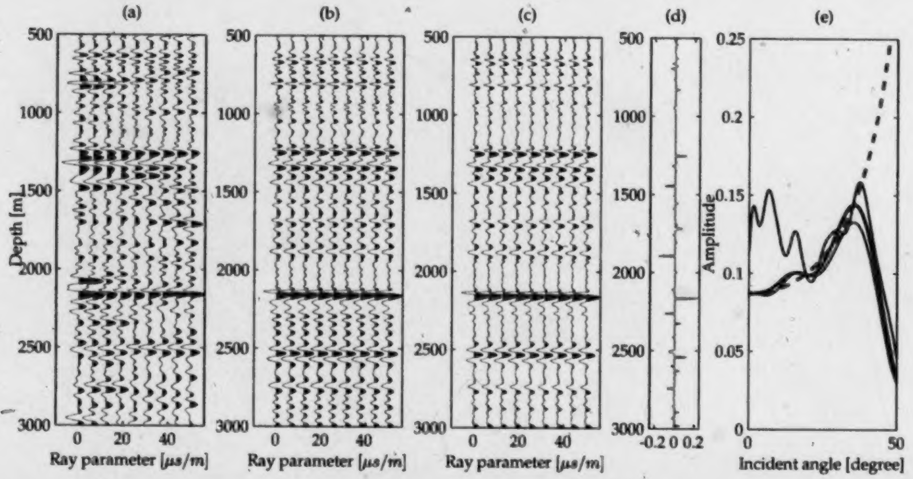


Figure 4.40: Zoomed-in common image gathers (CIGs) and AVA curves at $x = 7500$ m for the Marmousi data. (a) Migration. (b) PLSM. (c) SLSM. (d) Zero-offset reflectivity from the density and velocity model. (e) AVA curves for the event at depth 800 m.Red dashed: the theoretical curve. Green: migration. Blue: PLSM. Black: SLSM.

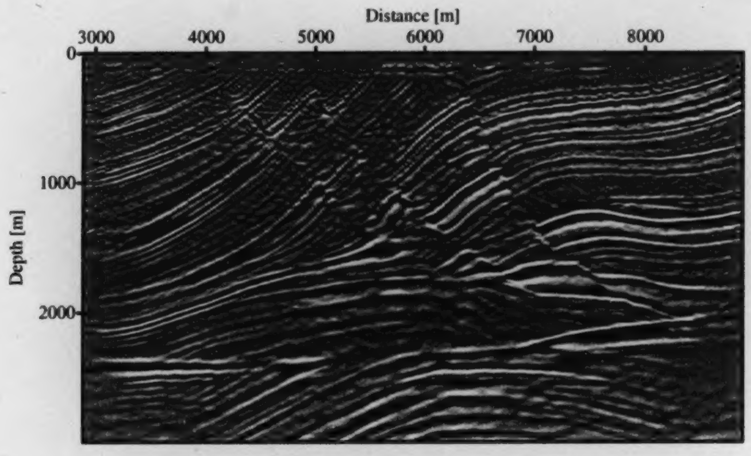


Figure 4.41: Stacked image by migration of the incomplete Marmousi data. 70% traces were randomly decimated from the original data.

Figure 4.41 shows the stacked images acquired by the conventional migration. It can be seen that the data missing has introduced some high-frequency artifacts in the migrated image. The deeper events are less affected. Two reasons might account for that. First, high-frequency waves are more attenuated than low-frequency waves. Second, in a common midpoint gather of seismic data, shallower events usually have larger moveout, the difference of arrival time between near offset and far offset. As we know, more dipping events are more vulnerable to aliasing introduced by sparse sampling of the offsets. With the disturbance of imaging artifacts, the fault system is blurred. These artifacts are suppressed by PLSM and SLSM (see Figure 4.42 and Figure 4.43). The deep anticline is better focused. It is clear that SLSM provides a cleaner and sharper image than PLSM. However, since the source wavelet is not included in the inversion, I couldn't push the regularization too hard. For this reason, I do not expect SLSM to give a very spiky image.

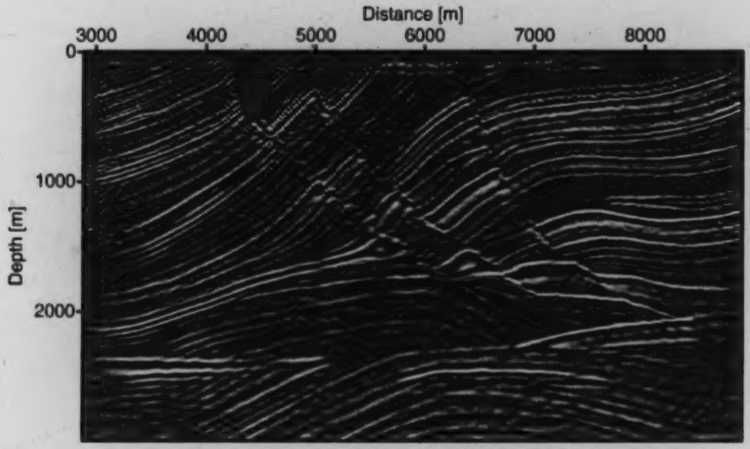


Figure 4.42: Stacked image by the preconditioned least-squares migration of the incomplete Marmousi data. 70% traces were randomly decimated from the original data.

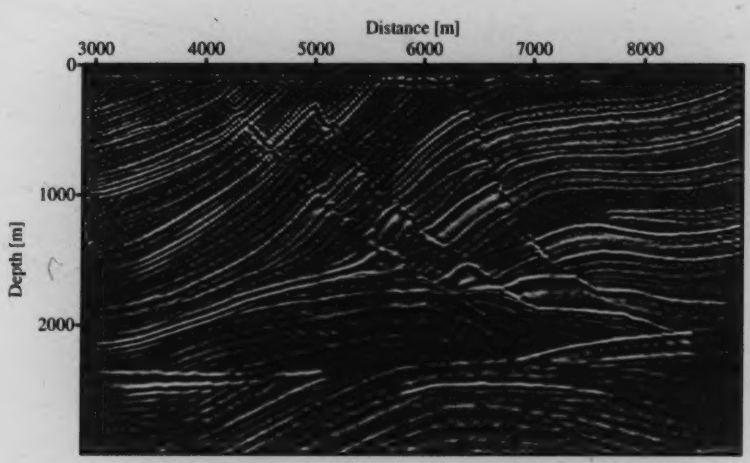


Figure 4.43: Stacked image by the sparse least-squares migration of the incomplete Marmousi data. 70% traces were randomly decimated from the original data.

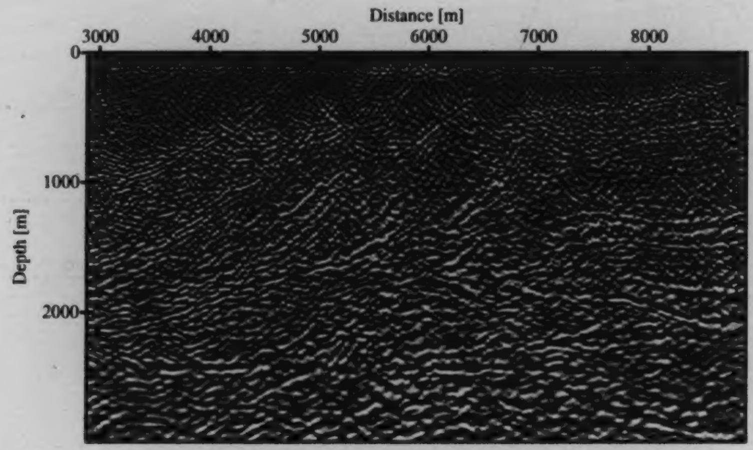


Figure 4.44: Image of constant ray parameter ($p_{h_0} = 0\mu s/m$) by migration of the incomplete Marmousi data. 70% traces were randomly decimated from the original data.

As a bonus of least-squares migration, I found that the inversion can fill in the near offset gaps of the seismic survey. Note that it is difficult for the conventional migration to acquire near offset information since even the complete data are lack of near offset traces (the minimum offset of the Marmousi data is 200 m). In real seismic surveys, this is very common because we usually do not put receivers too close to the source for security of the instrument. Therefore a straight-forward thought is that the migration will poorly illuminate the image of low ray parameters. On the other hand, least-squares migration takes advantage of the physical model (subvolume velocity information) and the regularization of the model to interpolate and extrapolate the wavefield. These naive ideas are supported by a comparison of the resolved image of a constant ray parameter, $p_{h_x} = 0\mu s/m$ (see Figure 4.44, Figure 4.45 and Figure 4.46). In the migration result, it is difficult to recognize the geological structure, and the image is seriously aliased. On the contrast, the images provided by PLSM and SLSM fundamentally improve the solution.

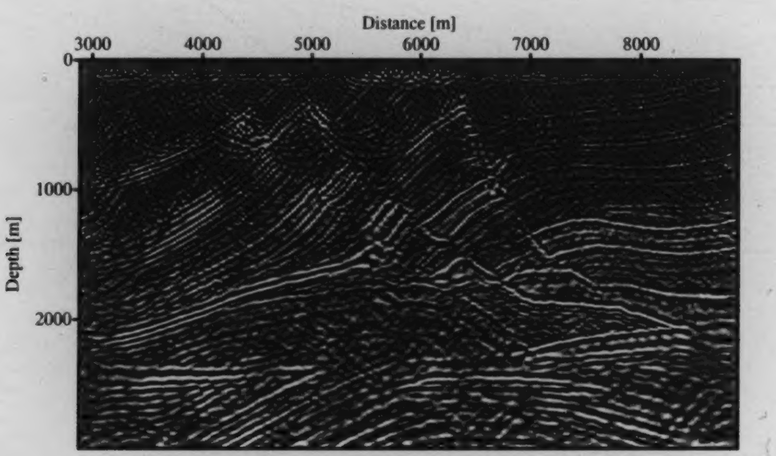


Figure 4.45: Image of constant ray parameter ($p_{h_0} = 0\mu s/m$) by the preconditioned least-squares migration of the incomplete Marmousi data. 70% traces were randomly decimated from the original data.

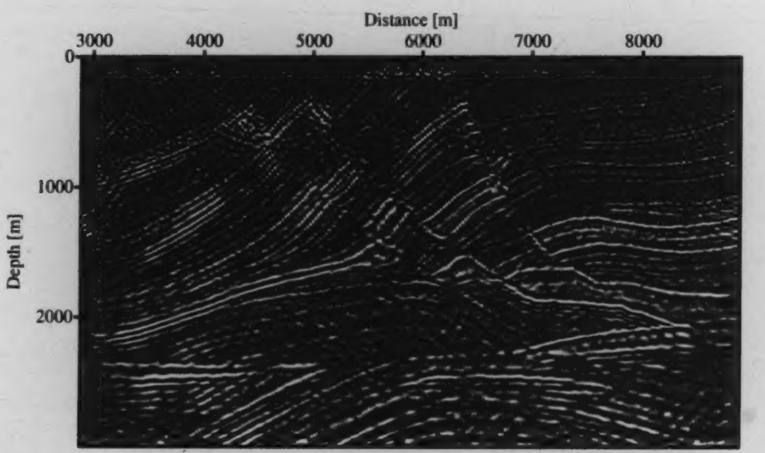


Figure 4.46: Image of constant ray parameter ($p_{h_0} = 0\mu s/m$) by the sparse least-squares migration of the incomplete Marmousi data. 70% traces were randomly decimated from the original data.

4.3 Summary

I have conducted extensive tests on the wave equation based methods of migration /inversion. As the examples have shown, the operator can efficiently handle complex geological settings and models with strong velocity variations. Regularized least-squares migration can remove the artifacts due to data acquisition footprint and operator mismatch. The computational cost of the inversion can be greatly decreased by a preconditioning strategy.

The combination of smoothing regularization and sparse regularization is robust at reconstructing high-frequency information about the earth model. The implemented inversion improves the image quality in both the ray parameter direction and the depth direction. However we may have two difficulties when we apply this method. First we should find a proper trade-off parameter to avoid over regularization, which may result in loss of valuable information. Second, very sparse solution can only acquired with the good knowledge of the wavelet. This is a difficult problem especially for the real data.

Chapter 5

Field data example

5.1 Introduction

In Chapter 4, I have compared the wave-equation migration and the least-squares migration with various synthetic datasets. As the results have shown, least-squares migration is more robust than the conventional migration in providing accurate structure and amplitude information. The nice point of synthetic data is that we know the model originating the data. Therefore the synthetic data allow us to conduct fair and well controlled comparison between the inverted model and the real model. On the other hand, field data are more complicated. First we do not know the real model. Second, the modelling mechanism is much more complicated than the formulas we have used. Third, the noise behavior could be unpredictable. Fortunately, we can utilize well log data, with higher resolution than seismic reflection data, to control the uncertainties of seismic processing and interpretation. In this chapter, I first evaluate the performance of various migration/inversion methods that has discussed in previous chapters with a field dataset from the Western Canadian Sedimentary Basin (WCSB). Then the accuracy of the inversion is evaluated by the synthetics from well log data. This is an important step toward the application of 3-D least-squares migration..

By far, in the geophysical community most case studies of AVO/AVA analysis have been based on the data preprocessed by NMO/DMO methods or the approximate model obtained with migration, the adjoint of the modelling. The result can be unsatisfactory due to model complexity and data loss. On the contrast, the regularized migration/inversion proposed in this thesis has the potential to provide coherent and accurate structure and amplitude for petrophysical feature extractions. The tests bellow are not a fully detailed case history but an evaluation of the algorithm with the support of well log data. To my knowledge, this is the first time this type of migration/inversion methods are calibrated with and tested against well log data.

5.2 Erskine data from WCSB

The data were gathered near a small town, Erskine, Southern Alberta, Canada. An orthogonal acquisition geometry was adopted to cover the survey area. 'Orthogonal' means the source lines are normal to the receiver lines. The small 3-D survey targets the Leduc reef, a carbonate play in the Western Canadian Sedimentary Basin (WCSB). The data were first binned, and a constant common-azimuth subset was extracted. The binned data consist of 157 inlines and 40 crosslines. The offset ranges from 75 m to 3000 m, with a highly uneven and sparse distribution (Figure 5.1). Each trace of the Erskine dataset contains 1000 samples with 2 ms interval. The inline CMP (common midpoint

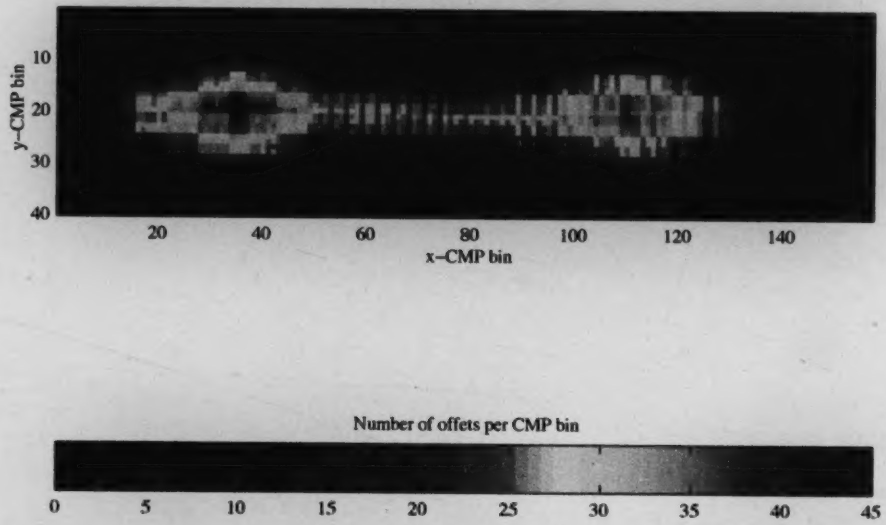


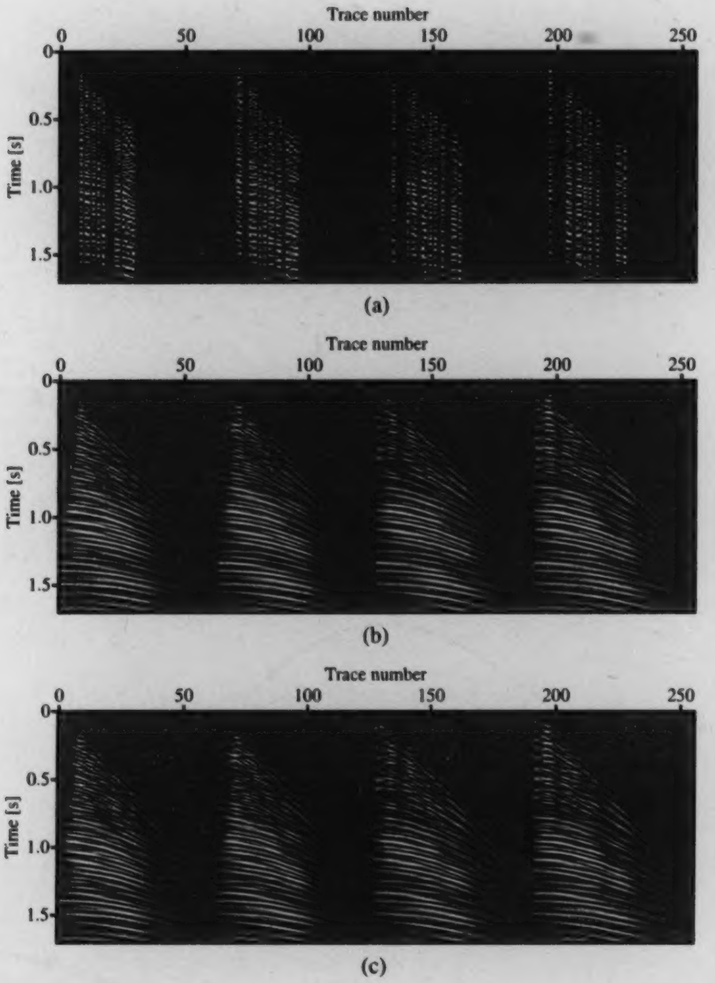
Figure 5.1: Offset distribution per CMP bin for the Erskine dataset (WCSB). The number of offsets in each bin is color coded.

gather) spacing is 33.5 m, and the cross line spacing is 50.29 m. The azimuth direction is NE 180.58786°. Some preprocessings, including spherical divergence, deconvolution, surface consistent statics and a band-pass filtering (10/15 Hz-65/80 Hz), have been conducted before applying the following tests of migration/inversion.

This particular dataset was also used for studies of data regularization by (Liu and Sacchi, 2004). These authors have analyzed the impact of data regularization, prior to wave-equation migration, on the quality of angle-domain common image gathers. Both interpolation and least-squares migration can address the problem of data incompleteness. However it is clear that least-squares migration should promise more than interpolation since it inverts not only the data sampling matrix but also the wave equation based modeling kernel. Detailed comparison of interpolation and least-squares migration will be considered in the near future.

5.3 Migration vs. LSM with smooth regularization

The Erskine dataset was used to compare the performance of migration, RLSM, and PLSM. As shown in the synthetic examples, the smooth regularization, adopted by RLSM and PLSM, is very efficient at reconstructing the data. It is important to test the validness of this technique in handling a sparse 3-D field data like Erskine dataset. Figure 5.2a shows 4 adjacent CMP gathers extracted from inline No. 10. Forward modeling, after inversion, is used to recover the data on the complete input grid. The resulting reconstructed gathers are depicted in Figures 5.2b (RLSM) and 5.2c (preconditioned RLSM). It is clear that both methods help to fill in the gaps between live traces. Furthermore, missing information in near offset and far offset is partially retrieved. Figures 5.3 shows the CIGs with offset ray-parameters ranging from 0-500 $\mu s/m$, with an interval of 6.25 $\mu s/m$ at inline No. 71, and crossline No. 10. Figure 5.3a portrays the migrated CIG with obvious aliasing artifacts and Figures 5.3b and 5.3c depict the least-squares inverted CIG after 4 and 11 iterations, respectively. Figure 5.3d shows the result of preconditioned RLSM after only 4 iterations with a quality similar to that of Figure 5.3c.



(a)

(b)

(c)

Figure 5.2: Comparison between observed and reconstructed data at inline No.10. (a) Original CMP gathers. (b) Reconstructed CMP gathers after 11 CG iterations. (c) Reconstructed CMP gathers after 4 preconditioned CG iterations.

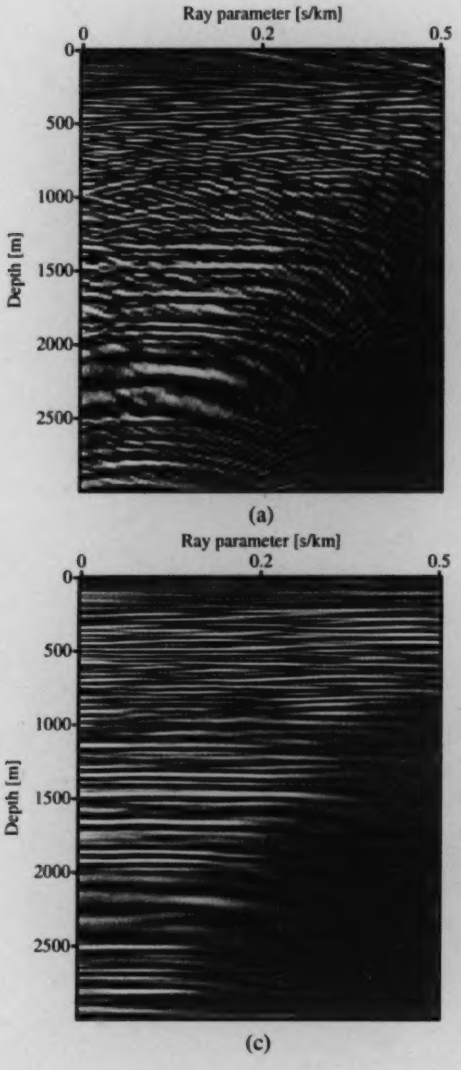


Figure 5.3: Common image gathers at crossline No. 10, inline No. 71. (a) CIG produced by migration. (b) CIG after 4 CG iterations. (c) CIG after 11 CG iterations. (d) CIG after 4 preconditioned CG iterations.

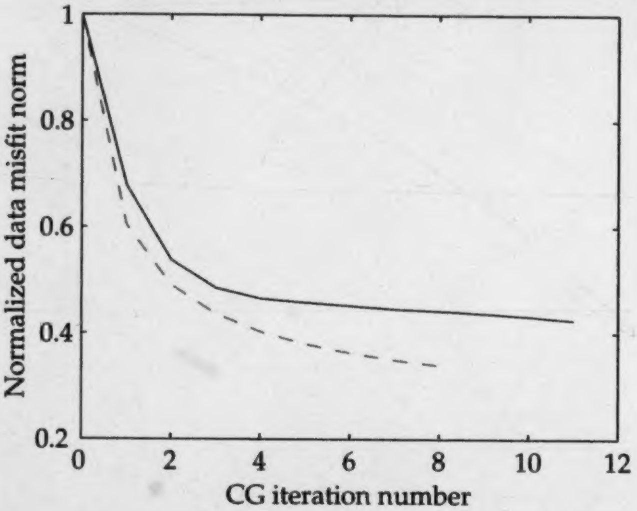


Figure 5.4: Normalized residual norm $||\mathbf{W}(\mathbf{Lm}-\mathbf{d})||^2$ versus CG iteration. Red dashed: preconditioned CG. Solid: CG.

The evolution of the normalized residual norm, $||\mathbf{W}(\mathbf{Lm}-\mathbf{d})||^2$, versus CG iteration can be seen in Figure 5.4. As expected, the preconditioned RLSM converges faster than RLSM with no preconditioning.

In general, in order to limit the computational cost of the inversion, the CG algorithm is stopped before complete convergence. In this case, I monitored a small subset of the reconstructed data (see Figures 5.2a, 5.2b and 5.2c) to control the degree of fitting. This is important since over-fitting leads to the formation of artifacts. Conversely, under-fitting leads to incomplete recovery of the missing observations.

The structural image is computed by stacking the CIGs along the offset ray-parameter. The stacked images obtained with migration, the regularized least-squares migration (RLSM) and the preconditioned regularized least-squares migration (PLSM) are displayed in Figure 5.5a, b and c, respectively. Both RLSM and preconditioned RLSM lead to better reflector continuity in low fold areas.

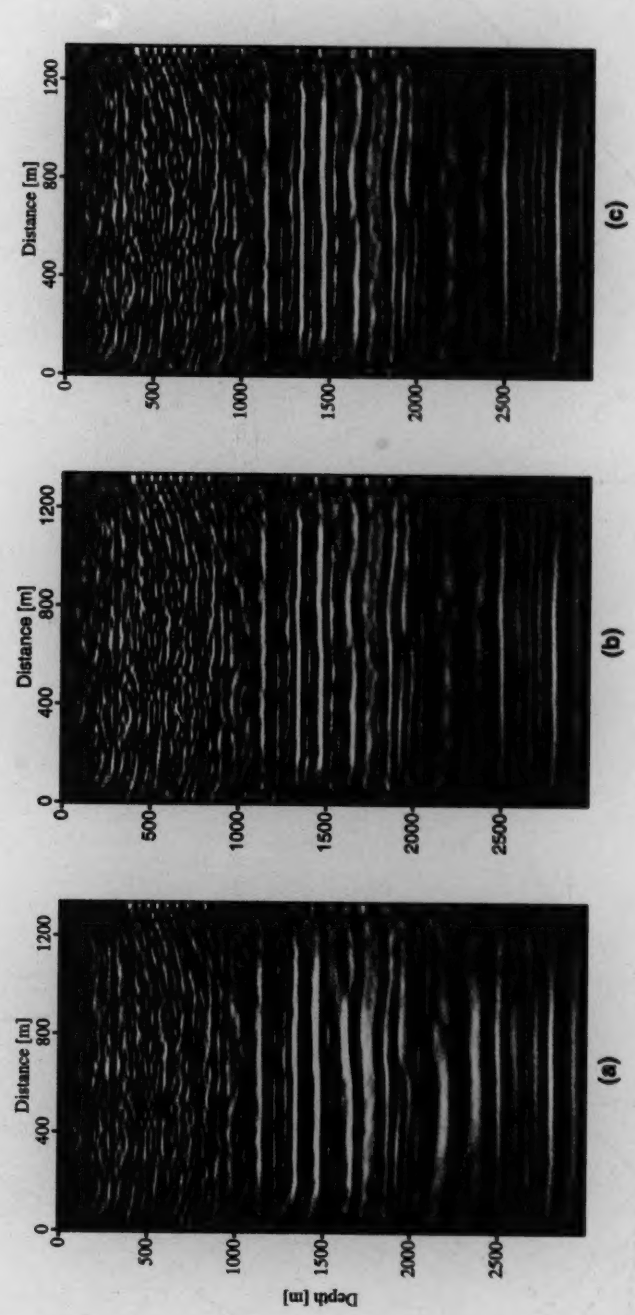


Figure 5.5: Common image gathers at crossline No. 10, inline No. 71. (a) CIG produced by migration. (b) CIG after 4 CG iterations. (c) CIG after 11 CG iterations.(d) CIG after 4 preconditioned CG iterations.

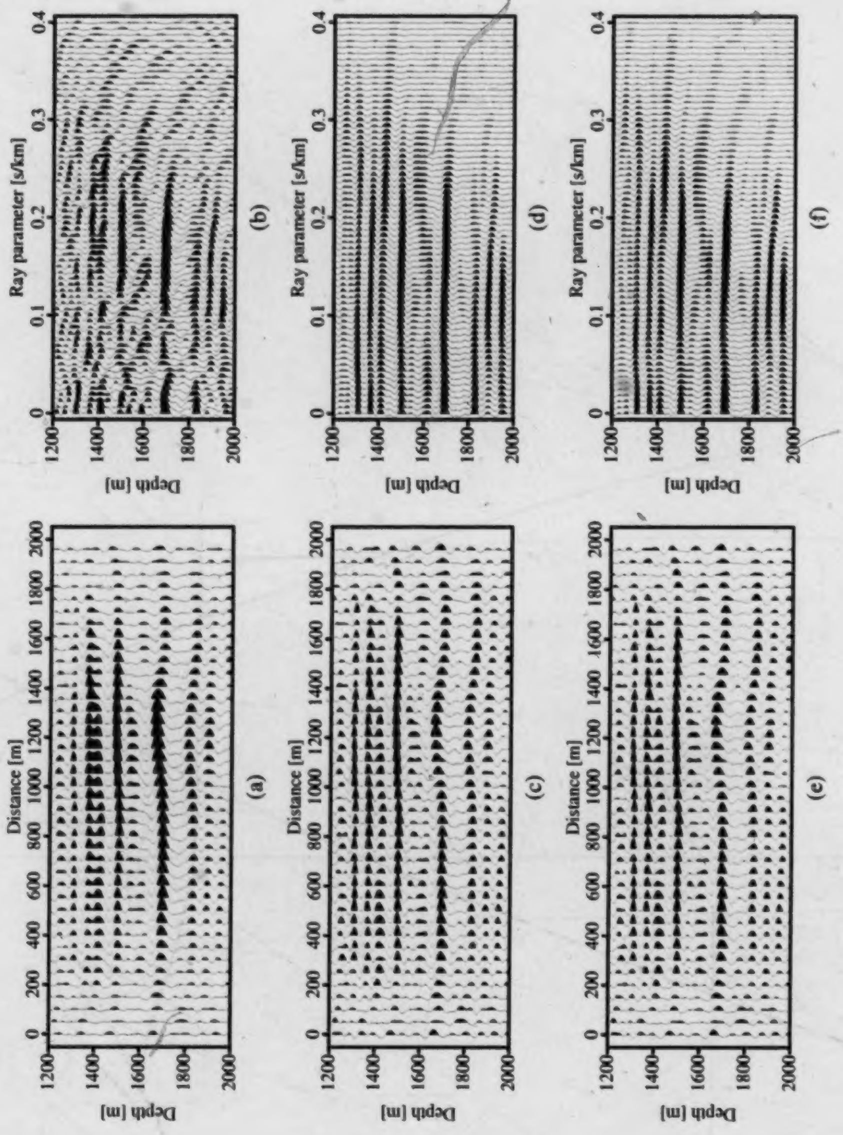


Figure 5.6: Detailed image of inline No.71 with CIG at crossline No. 10. (a) Stack after migration. (b) CIG corresponding to (a). (c) Stack after 11 CG iterations. (d) CIG corresponding to (c). (e) Stack after 4 preconditioned CG iterations. (f) CIG corresponding to (e).

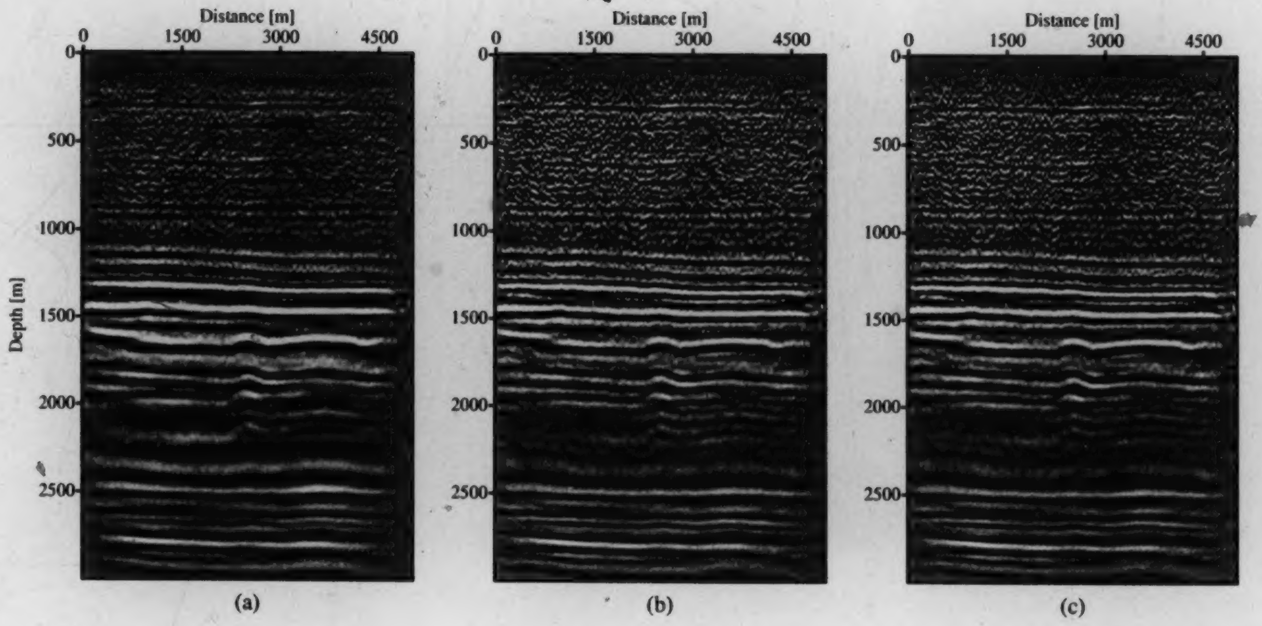


Figure 5.7: Stacked images of crossline No. 24. (a) Migration. (b) Least-squares migration after 11 CG iterations. (c) Preconditioned least-squares migration after 4 iterations.

Figure 5.6a, c and e show details of the stacked images displayed in Figure 5.5 and their associated CIGs (5.6b, d and f) at crossline No. 10. Note the considerably improved resolution. This effect can be explained as follows: First, by imposing smoothness on the inverted CIG, individual traces stack more coherently. In particular, part of the smearing produced by the aperture limitation (non-flatness at high ray-parameters) is attenuated and, therefore, the stacked CIG better preserves the high frequencies. Second, as mentioned earlier, least-squares migration automatically accounts for migration deconvolution described by Hu et al. (2001), which helps to sharpen the image. This is an important concept that can lead to higher resolution. As shown later, the addition of a vertical sparseness constraint can further increase the vertical resolution. The deconvolution effect is also visible in Figure 5.7 where I compare stacked images at crossline No. 24.

For better performance assessment, I generated a synthetic CIG based on available sonic log data from a well located at inline No. 76 and crossline No. 24. A second well, located outside the survey area, had density and sonic log information. Correlation of the two sonic logs allowed us to match the density log to the log located within the survey. This relatively crude approach appears to be justified, since the sonic logs agree very well (Figure 5.8).

The shear wave velocities are assumed to follow Castagna's mud-rock regression $V_s = (V_p - 1360m/s)/1.16$ (Castagna et al., 1985). Since Castagna's formula is not valid for carbonates, I restrict the AVA analysis to the Ellerslie (sandstone) and Banff formation (shale) (see Figure 5.9) as indicated in the stratigraphic column (Mossop and Shetsen, 1994). Unfortunately, by disregarding the deeper carbonates, I exclude the Leduc reef which is the actual exploration target. However, this way I do not introduce further uncertainties by attempting to estimate the carbonate shear wave velocities. With the compressional wave velocities, estimated shear wave velocities and the calibrated densities, I calculate the angle-dependent reflectivity traces using the Aki and Richards approximation of Zoeppritz's equations (Aki and Richards, 1980).

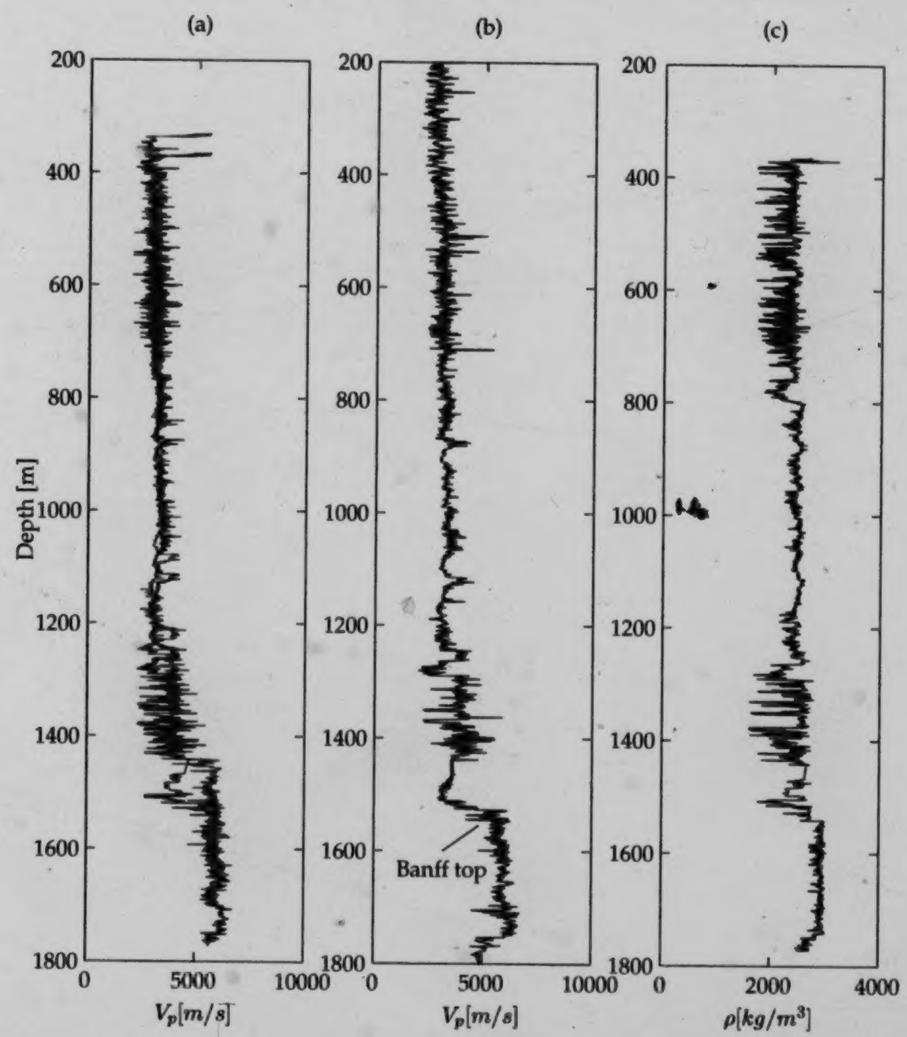


Figure 5.8: Depth correlation of log traces. (a) Sonic log trace (V_p) outside the survey area. The red curve is the trace at the original depth. The blue curve is the same trace with depth calibrated to the local trace shown in (b). (b) Sonic log trace at inline No. 76, crossline No. 24. (c) Density log outside the survey area. The depth has been adjusted to match the sonic log trace within the survey area.

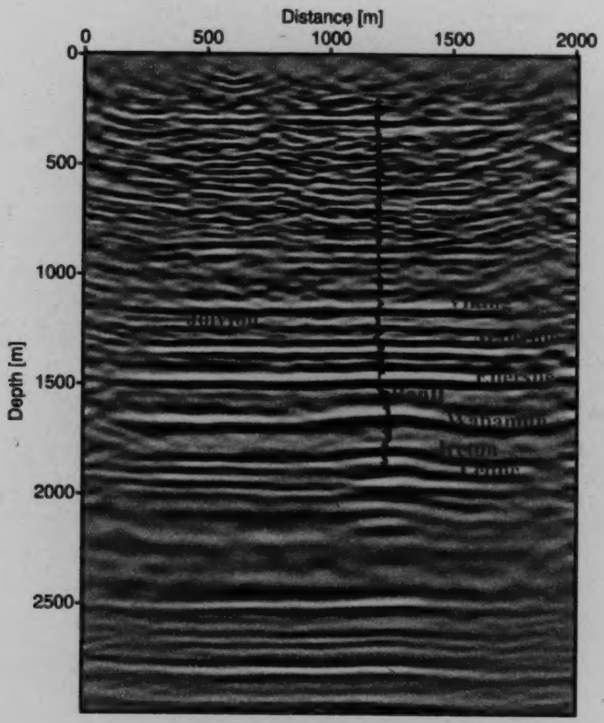


Figure 5.9: Strata correlation for the survey area of the Erskine dataset. The background is the stack of inline No. 76 and the blue curve is the local sonic log (same as Figure 5.8b). The stack and the sonic log correlate relatively well.

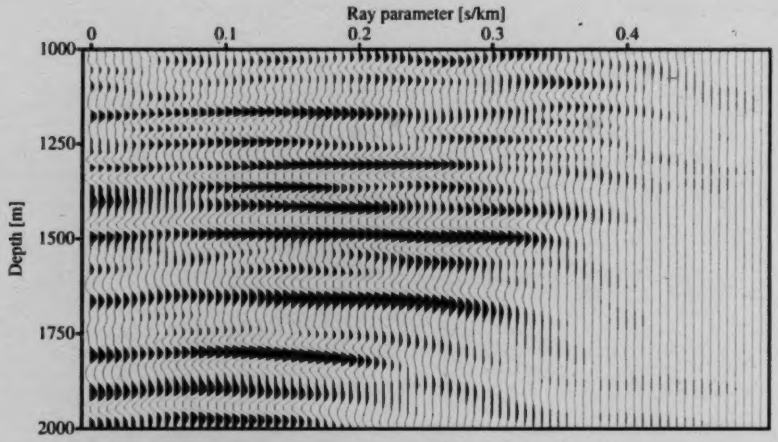


Figure 5.10: CIG at inline No. 76, crossline No. 24. See Figure 5.11b for the AVA and time conversion of the highlighted area.

The inverted CIG at the well location is displayed in Figure 5.10. Figure 5.11 compares the time converted Aki and Richards synthetic CIG and the inverted and time converted CIG in more detail. Given the degree of uncertainty associated with the synthetic and the sparseness of the field data, the match is acceptable, despite some discrepancies. The picked AVA for the prominent Ellerslie/Banff event at $0.7s$ in Figure 5.12 fits the synthetic AVA well within 12 to 27 degrees. Outside this angle range, it is difficult to acquire reliable amplitude due to the lack of data support.

5.4 PLSM vs. SLSM

As an experiment, I also processed the field data with the sparse least-squares migration (SLSM). To save computational time and avoid overfitting of the noisy data, I only ran the IRLS program for three iterations, each of which involved 5 iterations of preconditioned least-squares migration (PLSM).

Similar to what the test of the Marmousi data has shown, the sparse regularization suppresses the spurious imaging artifacts efficiently (see Figure 5.13). On the other hand, coherent signal is enforced. It is clear that the seismic events obtained with SLSM look cleaner and thinner than those with PLSM. Overlaping events are better resolved and

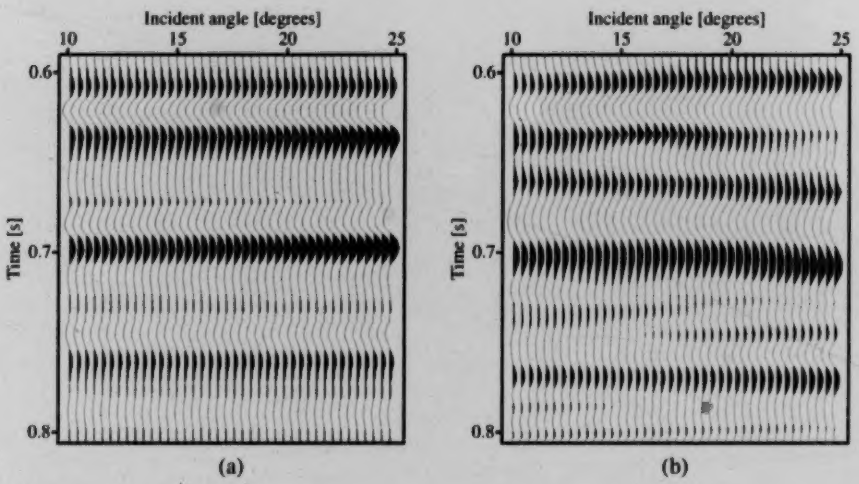


Figure 5.11: Comparison between the synthetic CIG and the inverted CIG. (a) Synthetic CIG. (b) Inverted CIG. Both CIGs are displayed in time domain. AVA curves for the event at $0.7s$ are shown in Figure 5.10. Since the events are quite flat, AVA is calculated from AVP using equation (A-5) with a zero inline dip angle.

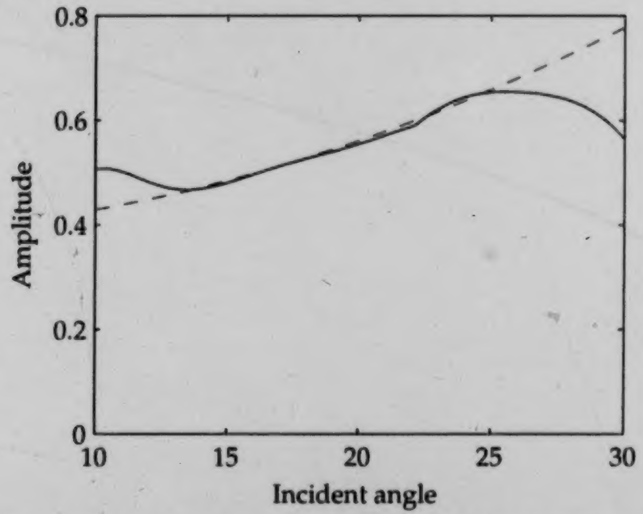
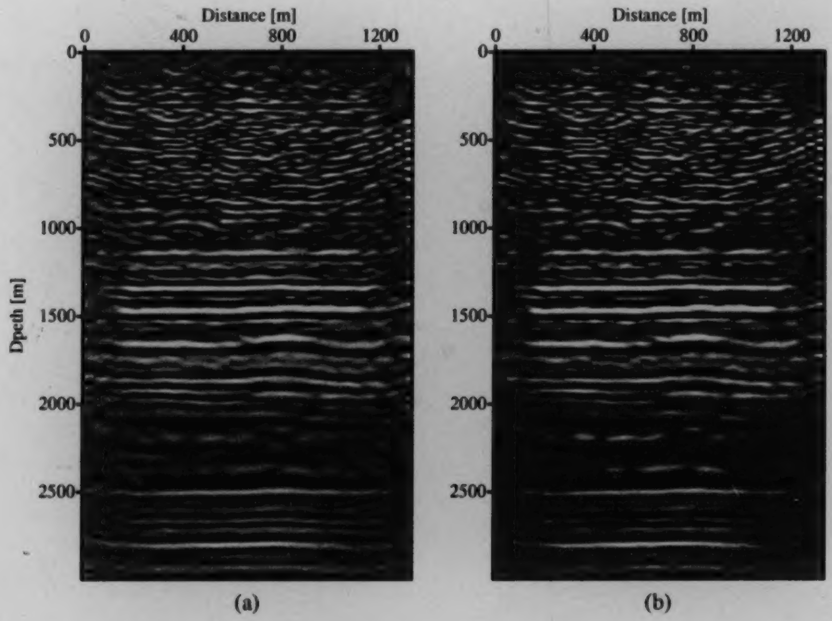


Figure 5.12: Comparison between the synthetic AVA and the inverted AVA. Dashed: synthetic AVA for the event at depth 1500 m. Solid: inverted AVA for the same event (see Figure refsynthetic:vs:inverted)

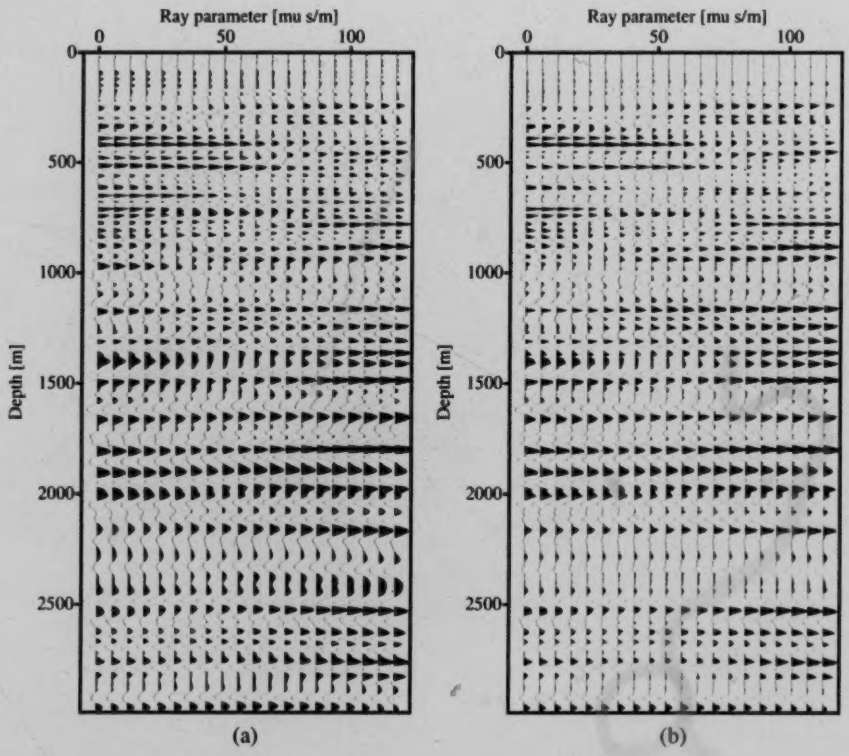


(a) (b)

Figure 5.13: Stacked images obtained by the preconditioned least-squares migration (PLSM) and the sparse least-squares migration (SLSM) (a) PLSM. (b) SLSM.

separated from each other. Figure 5.14 compares the common image gathers at inline no. 76 and crossline no. 24. Note that the overlapping events at depth in the result of PLSM are nicely separated by SLSM.

However, there are some traps in the application of SLSM processing. As shown in the stacked images, some original discontinuities in the result of PLSM, which display smearing and weak amplitude, are decimated together with the imaging artifacts. Three reasons may account for such information loss. First, big gaps in the data introduce aliasing in the solution. Second, velocity errors lead to smearing of the image. Third, over regularization will remove events of small amplitude, which has been examined in Chapter 3.



(a)

(b)

Figure 5.14: Inverted common image gathers by the preconditioned least-squares migration (PLSM) and the sparse least-squares migration (SLSM). (a) PLSM. (b) SLSM.

5.5 Stacked cubes and depth slices

As an output of the 3-D migration/inversion, stacked cubes are often used to locate potential hydrocarbon traps in reference to typical favorable structures. Therefore the resolution enhancement of the stacked cube is beneficial to the structure analysis. Figure 5.15 displays the results calculated by four methods: migration, regularized least-squares migration (RLSM), preconditioned least-squares migration and sparse least-squares migration (SLSM). It is evident that the least-squares migration methods provide much higher resolution than the standard migration. Even the edge area with low data folds are highly resolved. SLSM provides the highest resolution among four imaging methods. Similar phenomena is observed in the comparison of the depth slices at $z = 1800$ m (see Figure 5.16). The image edge of the depth slice is best preserved by SLSM.

5.6 Summary

In this Chapter, I have used a field dataset to test the proposed least-squares migration algorithms. The tests further confirm that the regularized migration/inversion improves the coherence of the image gathers. More importantly, the amplitude accuracy is significantly improved with the inversion of the modelling kernels. This is supported by a side-by-side comparison of the inverted image gathers and the synthetic gathers with the well log data. The test of real data and verification by well log synthetics is an important step towards applications of these methods.

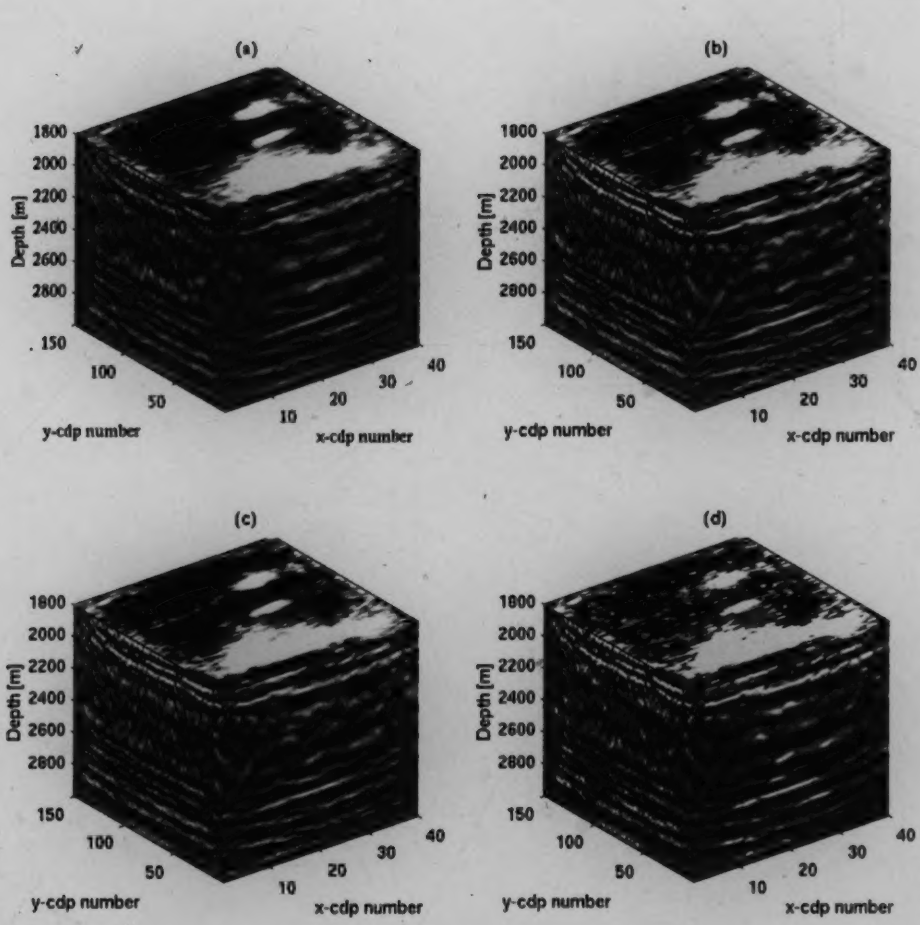

(a)
(b)
(c)
(d)
Depth [m]
1800
2000
2200
2400
2600
2800
150
100
50
y-cdp number
10
20
30
40
x-cdp number

Figure 5.15: Stacked cubes obtained by migration, regularized least-squares migration (RLSM), preconditioned least-squares migration (PLSM) and sparse least-squares migration (SLSM). (a) Migration. (b) RLSM (11 iterations) (c) PLSM (4 iterations) (d) SLSM. 3 iterations of the IRLS algorithm.

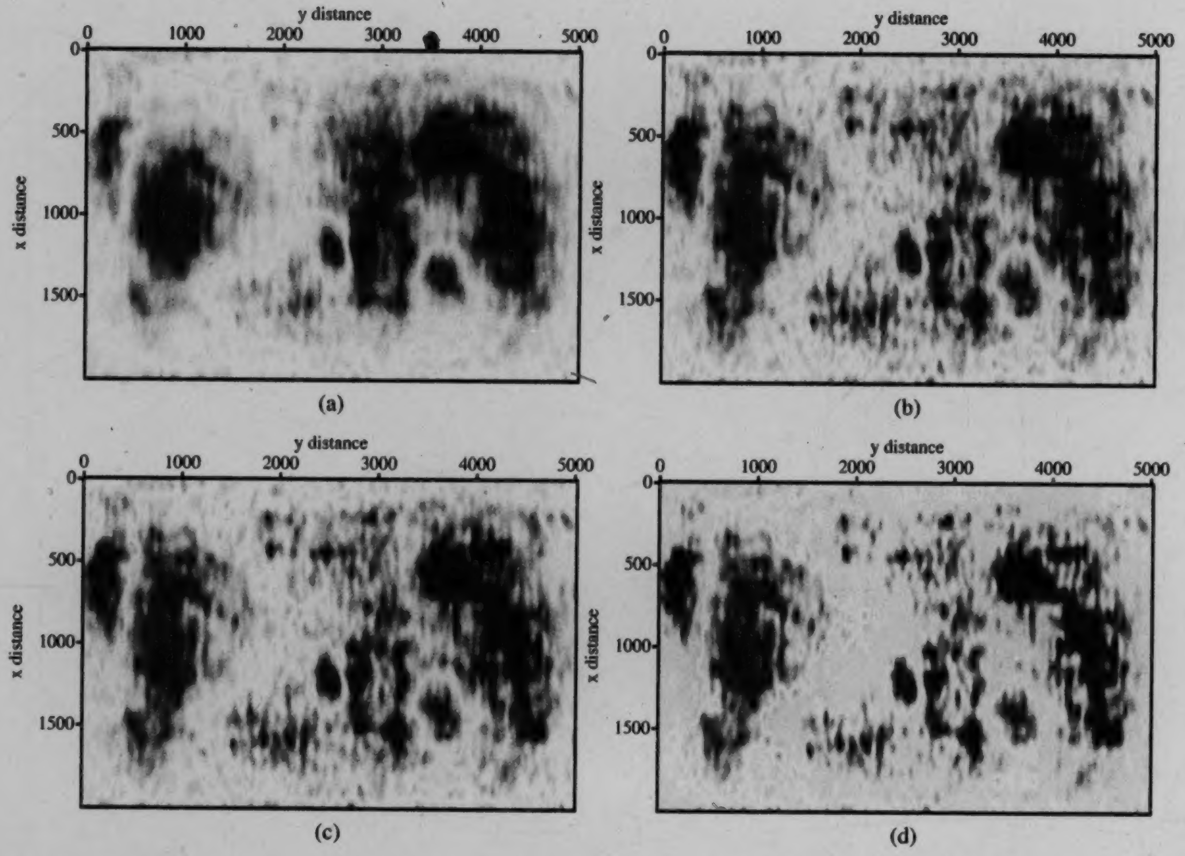


Figure 5.16: Stacked depth slices at $z = 1800$ m, obtained by migration, regularized least-squares migration (RLSM), preconditioned least-squares migration (PLSM) and sparse least-squares migration (SLSM). (a) Migration. (b) RLSM (11 iterations) (c) PLSM (4 iterations) (d) SLSM. 3 iterations of the IRLS algorithm.

Discussion and conclusions

The goals of seismic migration/inversion is two-fold. First, we would like to image highly resolved structures of the earth's interior. Second, we want to make full use of the data in the prestack volume to retrieve physical parameters (rock properties and/or hydrocarbon indicators) by using angle dependent reflectivity. However, seismic migration/inversion is difficult. The main difficulties lie in the complexity of the earth model, the unknown modeling mechanism and limited resolution of the recorded data. It is always challenging to find a precise representation of the earth. The earth is not acoustic and not even strictly elastic. But for practical reasons, we explore the earth based on elastic or acoustic wave theory. Especially, since the compressional reflection waves are usually stronger than shear waves, the acoustic theory proved to be a robust tool to study the subsurface. Even after assuming acoustic media, the problem is still complex due to the existence of multiples, which leads to a non-linear inverse problem. Furthermore, the recorded data always contain unpredictable noise, and we can only record data within a limited aperture. All these problems limit our ability to fully solve the inverse problem.

Geophysicists have provided various solutions to solve these problems. First before seismic migration/inversion, the data are preprocessed to remove multiples. As a result, the problem is changed into a linear problem. Second, various signal processing methods, for example deconvolution, coherent noise suppression, etc, are used to increase the signal-to-noise ratio. Furthermore, interpolation methods are adopted to treat the data aliasing problem.

In this thesis, I have proposed a strategy to deal with 3-D prestack data imaging using iterative inversion. I use an economical 3-D wave equation based operator, common-azimuth seismic modelling/migration that can handle real data size situations. Implemented with the split-step and split-step PSPI corrections, the operator can be used to image complex structures.

As a benefit, *a priori* information about the model is encoded in the inversion to regularize the solution. Two kinds of regularization have been adopted to contribute to the resolution enhancement. First, I force smoothness of the amplitude in the inline offset ray-parameter direction, which abounds in coherence of common image gathers. Second, I suppress the wavelet effects in the depth direction by minimizing Cauchy norm of the stacked image. The field data test supports the assertion that regularized least-squares migration provides improved amplitude robustness and image resolution even when the data are very sparse and aliasing hampers the AVA analysis.

Good quality CIGs with accurate amplitude are key for amplitude supported seismic interpretation that aims at rock and fluid properties estimation. I have conducted extensive tests to show that the regularized least-squares migration can properly preserve the amplitude even when data aliasing and noise is present. More tests on 3-D data with complex structure might be necessary to further confirm the accuracy of common-azimuth operator. For example, the 3-D Marmousi data are suitable both for structure and amplitude studies.

Care should be taken when we apply the sparse least-squares migration. Some tests have shown that over emphasis of sparseness leads to loss of valuable information that is often contained in events with small amplitudes. This problem is also encountered in techniques for post stack inversion of seismic data that are based on the sparse reflectivity assumption. To obtain very sparse solutions, a quite accurate model of the wavelet is required.

The iterative inversion is implemented with preconditioned conjugate gradients (CG) methods. Preconditioning is utilized to cut computational costs. In my tests of quadratically regularized least-squares migration, good convergence can be achieved within 4 iterations or so of preconditioned RLSM, which is a cost equivalent to 8 conventional migrations. This is a dramatic improvement over regularized least-squares migration with a cost on the order of 20 migrations. The sparse least-squares migration is more expensive. Usually it takes 3-4 iterations of iterative reweighted least-squares (IRLS) algorithm, and each IRLS iteration needs convergence of the CG algorithm. Therefore, at this stage, perhaps this method is currently more suitable for 2-D applications or for target oriented studies of 3-D data sets.

A potential application of this technique is to acquire high resolution azimuthal common image gathers (land data) . We can apply the algorithm to different azimuths. Given proper azimuthal velocity models, we can invert for azimuthal amplitude variations. The latter can be used to measure fracture density and fracture strike within carbonates (Rüger, 1996; Alhavas et al., 2003; Gray and Head, 2000). Furthermore, the technique can be used for carbon dioxide sequestration studies. In recent years, new efficient ways for carbon dioxide disposal have been proposed. The gas can be injected to depleted oil and gas reservoirs (Holloway et al., 1996; Wong et al., 1999). It is clear that detail information about the fracture distribution of the subsurface is of chief importance when planning a gas injection operation. Time-lapse analysis of multi-azimuth data can be used to help to monitor the disposal site.

Clearly, there exist computationally less demanding techniques, like interpolation prior to migration, that attempt to address some of the data issues discussed in this dissertation. The amplitude is then corrected by the inverse of imaging Jacobian during migration. Hence, to better appreciate the added benefits of least-squares migration, a careful comparative study between more conventional processing and least-squares migration is in order.

In spite of this caveat, least-squares migration deserves special attention as it is the unifying link between imaging and inversion. It allows us to fit the seismic data and, at the same time, impose geophysically sensible constraints on the seismic model.

The main contributions of this thesis can be summarized as follows:

- Following ideas proposed in the literature (Nemeth et al., 1999; Kuehl and Sacchi, 1999; Prucha and Biondi, 2002), the thesis extended 2-D least-squares AVP/AVA migration to the 3-D case. An efficient common-azimuth operator was adopted to process 3-D seismic data in a computationally feasible manner

- I have introduced a sparseness constraint to enhance the resolution of seismic events in the inverted CIGs.

- I calibrated the algorithm with synthetics obtained via well log data. This appears to be the first serious attempt to validate true amplitude AVA responses against well log data.

I expect that other regularization schemes will be proposed in the future. For instance, inversion of CIGs that are parameterized in terms of AVA intercept and gradient combined with a spatial smoothness constraint that conforms to the geological structure may be one avenue to further develop least-squares migration. In this case, the problem becomes over-determined with a considerable reduction of model parameters. Steps in this direction were started by Feng (2004), using kirchhoff operators. The idea can also be adopted with wave-equation operators and iterative inversion.

References

Akaike, H., 1974, A new look at the statistical model identification: IEEE Transactions on Automatica Control, **19**, 716–723.

Aki, K., and Richards, P. G., 1980, Quantitative seismology theory and methods: W. H. Freeman and Company.

Alhavas, K., Ameen, M., Nebrija, E., and MacBeth, C., 2003, Delineating of fracture anisotropy signatures in wudayhi field by azimuthal seismic data: The Leading Edge, **22**, 1202–1211.

Aminzadeh, F., Burkhard, N., Rocca, F., and Wyatt, K., 1994, SEG/EAGE 3-D modeling project: 2nd update: The Leading Edge, **13**, 949–952.

Arfken, G., 1985, Mathematical methods for physicists: Academic Press.

Bagaini, C., Burkhard, N., Rocca, F., and Wyatt, K., 1995, Data parallel implementation of 3-D PSPI: 65th Ann. Internat. Mtg., Soc. of Expl. Geophys., Expaneded Abstracts, pages 188–191.

Berkhout, A. J., 1982, Seismic migration: Elsevier, Amserdam.

Berkhout, A. J., 1984, Seismic migration: Imaging of acoustic energy by wave field extrapolation (b. practical aspects): Elsevier.

Bevc, D., 1997, Imaging complex structure with semirecursive kirchhoff migration: Geophysics, **62**, 577–588.

Beylkin, G., and Burridge, R., 1990, Linearized inverse scattering problems in acoustics and elasticity: Wave motions, **12**, 15–52.

Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform: J. Math. Phys., **26**, 99–108.

Biondi, B., and Palacharla, G., 1996, 3-D prestack migration of common-azimuth data: Geophysics, **61**, no. 6, 1820–1832.

Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: Geophysics, **63**, 574–588.

Biondi, B., 2002, Stable wide-angle fourier finite-difference downward extrapolation of 3-D wavefields: Geophysics, **67**, 872–882.

Biondi, B., 2003a, 3-D seismic imaging: Standford University.

——— 2003b, Equivalence of source-receiver migration and shot-profile migration: Geophysics, **68**, no. 4, 1340–1347.

Black, J. L., Scheleicher, K. L., and Zhang, L., 1993, True-amplitude imaging and dip moveout: Geophysics, **58**, 47–66.

Bleistein, N., and Cohen, J. K., 1979, Direct inversion procedure for Claerbout's equations: Geophysics, **44**, no. 6, 1034–1040.

Bleistein, N., and Gray, S. H., 2002, A proposal for common-opening-angle migration/inversion: center for wave phenomena, Colorado School of Mines.

Bleistein, N., Cohen, J. K., and Stockwell Jr., J. W., 2000, Mathematics of multidemensional seismic imaging, migration, and inversion: Springer-Verlag.

Castagna, J. P., Batzle, M. L., and Eastwood, R. L., 1985, Relationship between compressional wave and shear wave velocities in clastic silicate rocks: Geophysics, **50**, 571–581.

Cerveny, V., 2000, Seismic ray theory: Cambridge University Press, Cambridge.

Charbonnier, P., Blanc-Feraud, L., Aubert, G., and Barlaud, M., 1997, Deterministic edge-preserving regularization in computed imaging: IEEE Transactions on Image Processing, **6**, 298–311.

Claerbout, J. F., 1971, Toward a unified theory of reflector maping: Geophysics, **36**, 467–481.

Claerbout, 1985, Imaging the earth's interior: Blackwell Science Inc.

Claerbout, J. F., 2004, Image estimation by example: Geophysical soundings image construction: multidimensional autoregression: Stanford Exploration Project. URL http://sepwww.stanford.edu/sep/prof/.

Clayton, R. W., and Stolt, R. H., 1981, A Born-WKBJ inversion method for acoustic reflection data: Geophysics, **46**, 1559–1567.

Cohen, J. K., and Bleistein, N., 1979, Velocity inversion procedure for acoustic waves: Geophysics, **44**, 1077–1087.

de Bruin, C. G. M., Wapenaar, C. P. A., and Berkhout, A. J., 1990, Angle-dependent reflectivity by means of prestack migration: Geophysics, **55**, 1223–1234.

Debeye, H. W. J., and van Riel, P., 1990, LP-norm deconvolution: Geophys. Prosp., **38**, 381–404.

Downton, J., and Lines, L., 2004, Three term AVO waveform inversion: CSEG Annual Convention, CDROM,5 pages.

Duquet, B., Marfurt, K. J., and Dellinger, J. A., 2000, Kirchhoff modeling, inversion for reflectivity, and subsurface illumination: Geophysics, **65**, 1195–1209.

Ehinger, A., and Lailly, P., 1991, Prestack linearized inversion with a prior information: a step beyond prestack depth migration: 61st Ann. Internat. Mtg., Soc. Expl. Geopl., Expanded Abstracts, pages 1289–1292.

Etgen, J. T., 1994, Stability of explicit depth extrapolation through laterally-varying media: 64th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 1266–1269.

Feng, J., 2004, Rock properties inversion with kirchhoff ava migration/inversion: MSc. thesis, the University of Alberta.

Fomel, S., and Claerbout, J. F., 2003, Multidimensional recursive filter preconditioning in geophysical estimation problems: Geophysics, **68**, 577–588.

Fomel, S., 2004, Theory of 3-D angle gathers in wave-equation imaging: 74th Annual Internat. Mtg., Soc. Expl. Geophys., SEG Expanded Abstracts, pages 1053–1056.

Gazdag, J., 1978, Wave equation migration with the phase shift method: Geophysics, **43**, 1342–1351.

Gazdag, J., 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, 124–131.

Gray, F. D., and Head, K., 2000, Fracture detection in the manderson field-a 3-D AVAZ case history: 62nd Mtg., EAGE, Session A0015.

Gray, S. H., Etgen, J., Dellinger, J., and Whitmore, D., 2001, Seismic migration problems and solutions: Geophysics, **66**, 1622–1640.

Hale, D., 1984, Dip-moveout by forier transform: Geophysics, **49**, 741–757.

Hanke, M., and Hansen, P. C., 1993, Regularization methods for large-scale problems: Survey on Mathematics for Industry, **49**, 253–315.

Hansen, P. C., 1998, Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion: SIAM monographs on mathematical modeling and computation.

Hestenes, M., and Steifel, E., 1952, Methods of conjugated gradients for solving linear systems: Nat. Bur. Standards J. Res., **49**, 403–436.

Hill, N. R., 1990, Gaussian beam migration: Geophysics, **55**, 1416–1428.

Hill, N. R., 2001, Prestack gaussian-beam depth migration: Geophysics, **66**, 1240–1250.

Holloway, S., Heederik, J. P., van der Meer, L. G. H., Czermichowski, L., Harrison, R., Lindeberg, E., Summerfield, I. R., Rochelle, C., Schwarzkopf, T., Kaarstad, O., and Berger, B., 1996, The underground disposal of carbon dioxide: British Geological Survey, Keyworth, Nottingham, U.K.

Hu, I., Schuster, G., and Valasek, P., 2001, Poststack migration deconvolution: Geophysics, **66**, 1528–1537.

Huang, L. J., and Wu, R., 1996, 3D prestack depth migration with acoustic pseudo-screen propagators: Mathematical Methods in Geophysical Imaging IV, SPIE, **2822**, 40–51.

Huang, L. J., Sun, H., Fehler, M. C., and Li, Z., 2003, Controlled-aperture wave-equation migration: 73rd Annual Mtg., Soc. Expl. Geopl., Expanded Abstracts, pages 1087–1090.

Jin, S., Madariaga, R., virieux, J., and Lambare, G., 1992, Two-dimensional asymptotic iterative elastic inversion: Geophys. J. Int., **108**, 575–588.

Jin, S., Mosher, C. C., and Wu, R. S., 2002, Offset-domain pseudoscreen prestack depth migration: Geophysics, **67**, 1895–1902.

Kessinger, W., 1992, Extended split-step fourier migration: SEG Expanded Abstracts, pages 917–920.

Kostov, C., 1990, Multichannel seismic experiment with a drill-bit source: PhD thesis, Stanford University.

Kuehl, H., and Sacchi, M. D., 1999, Least-squares split-step migration using the hartley transform: 69th Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded Abstracts, pages 1548–1551.

Kuehl, H., and Sacchi, M. D., 2003, Least-squares wave-equation migration for AVP/AVA inversion: Geophysics, pages 262–273.

Kuehl, H., 2002, Least-squares wave-equation migration/inversion: PhD. thesis, University of Alberta.

Liner, C. L., 1990, General theory and comparative anatomy of dip moveout: Geophysics, **55**, 595–607.

Liu, B., and Sacchi, M., 2004, Mininum weighted norm interpolation of seismic records: Geophysics, **69**, 1560–1568.

Liu, B., Sacchi, M., and Kuehl, H., 2003, 2-D/3-D seismic wavefield reconstruction for AVA imaging: 73rd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 235-238.

Miller, D., Oristaglio, M. L., and Beylkin, G., 1987, A new slant on seisic imaging: Migration and integral geometry: Geophysics, **52**, 943–964.

Mosher, C. C., Keho, T. H., Weglein, A. B., and Foster, D. J., 1996, The impact of migration on AVO: Geophysics, **61**, 1603–1615.

Mossop, G., and Shetsen, I., 1994, Geological atlas of the western canada sedimentary basin: Canadian Society of Petroleum Geologists and Alberta Research Council, Calgary, Alberta, URL http://www.ags.gov.ab.ca /publications /ATLAS_WWW/ATLAS.shtml.

Mufti, I. R., Pita, J. A., and Huntley, R. W., 1996, Finite-difference depth migration of exploration-scale 3-D seismic data: Geophysics, **61**, 776–794.

Müler, G., 1971, Direct inversion of seismic observatios: J. of Geophysics, **37**, 225–235.

Nemeth, T., Wu, C., and Schuster, G. T., 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

Notfors, C. D., and Godfrey, R. J., 1987, Dip moveout in the frequency-wavenumber domain: Geophysics, **52**, 1718–1721.

O'Brien, M. J., and Gray, S. H., 1996, Can we image beneath salt: The Leading Edge, **15**, 17–22.

Oldenburg, D. W., Scheuer, T., and Levy, S., 1983, Recovery of the acoustic impedance from reflection seismograms: Geophysics, **48**, 1318–1337.

Popovici, A. M., 1996, Prestack migration by split-step DSR: Geophysics, **61**, no. 5, 1412–1416.

Prucha, M., and Biondi, B., 2002, Subsalt event regularization with steering filters: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 1176–1179.

Prucha, M., Biondi, B., and Symes, W., 1999, Angle-domain common image gathers by wave-equation migration: SEG Expanded Abstracts, pages 824–827.

Rickett, J. E., and Sava, P. C., 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889.

Ristow, D., and Rühl, T., 1994, Fourier finite-difference migration: Geophysics, **59**, 1882–1893.

Ronen, S., Nichols, D., Bale, R., and Ferber, R., 1995, Dealiasing DMO: Good-pass, bad-pass, and unconstrained: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 743–746.

Rüger, A., 1996, Reflection coefficients and azimuthal AVO analysis in anisotropic media: PhD thesis, Center for Wave Phenomena, Colorado School of Mines.

Saad, Y., 1991, Numerical methods for large eigenvalue problems: URL http://www-users.cs. umn.edu/ saad/books.html.

Sacchi, M. D., and Ulrych, T. J., 1995, High resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169–1177.

Sacchi, M. D., Constantinescu, C. M., and Feng, J., 2003, Enhancing resolution via non-quadratic regularization-next generation of imaging algorithms: CSEG Annual Convention, CDROM, 5 pages.

Sacchi, M. D., 1997, Reweighting strategies in seismic deconvolution: Geophysical Journal International, **129**, 651–656.

Sava, P., and Fomel, S., 2000, Angle-gathers by fourier transform: Stanford Exploration Project, **103**, 119–130.

Sava, P., and Fomel, S., 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Sava, P., Biondi, B., and Fomel, S., 2001a, Amplitude-preserved common image gathers by wave-equation migration: 71st Annual Mtg., Soc. Expl. Geoph., Expanded Abstracts, pages 296–299.

—— 2001b, Amplitude-preserved common image gathers by wave-equation migration: SEP Report 108, Stanford University.

Scales, J. A., and Smith, M., 1994, Introductory geophysical inverse theory: Samizdat Press.

Scales, J. A., 1987, Tomographic inversion via the conjugate gradient methods: Geophysics, **52**, 179–185.

Scales, J. A., 1997, Theory of seismic imaging: Samizdat Press.

SNL, 1998, SEG/EAGE salt C3 dataset: URL http: //www.cs.sandia.gov / ~ccober/seismic/ salt_c3.html.

Stoffa, P. L., Fokkema, J. T., de Luna Freire, R. M., and Kessinger, W. P., 1990, Split-step fourier migration: Geophysics, **55**, 410–421.

Stolt, R. H., and Bensen, A., 1986, Seismic migration - theory and practice: Geophysical Press, London - Amsterdam.

Stolt, R. H., 1978, Migration by fourier transform: Geophysics, **43**, 23–48.

Tarantola, A., 1987, Inverse problem theory-methods for data fitting and model parameter estimation: Elsevier.

Taylor, H. L., Banks, S. C., and McCoy, J. F., 1979, Deconvolution with the l_1 norm: Geophysics, **44**, 39–52.

Trad, D., Ulrych, T., and Sacchi, M. D., 2003, Latest views of the sparse radon transform: Geophysics, **68**, 386–399.

Tygel, M., and Hubral, P., 1984, Transient representation of the sommerfield-weyl integral with application to the point source response from a planar acoustic inteface: Geophysics, **49**, 1495–1505.

Ursin, B., 1983, Review of elastic and electromagetic wave propagation in horizontally layered media: Geophysics, **48**, 1063–1081.

Valenciano, A. A., and Biondi, B., 2003, 2-D deconvolution imaging condition for shot-profile migration: SEG Expanded Abstracts, pages 1059–1062.

Versteeg, R., 1994, The marmousi experience: velocity model deternation on a synthetic complex data set: The Leading Edge, **13**, 927–936.

Wang, J., Kuehl, H., and Sacchi, M. D., 2004, High resolution least-squares wave equation AVA imaging: Feasibility study with a data set from the western canadian sedimentary basin: 74th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 1101–1104.

Wang, J., Kuehl, H., and Sacchi, M. D., 2005, High resolution wave equation AVA imaging: Algorithm and tests with a data set from the western canadian sedimentary basin: Geophysics, **70**, S91–S99.

Wapenaar, C. P. A., and Berkhout, A. J., 1985, Wave field extrapolation techniques for inhomogeneous media which include critical angle events. part I: methods using the one-way wave equations: Geophysical Prospecting, **33**, 1138–1159.

Wapenaar, C. P. A., and Berkhout, A. J., 1986a, Wave field extrapolation techniques for inhomogeneous media which include critical angle events. part II: methods using the two-way wave equations: Geophysical Prospecting, **34**, 147–1179.

——— 1986b, Wave field extrapolation techniques for inhomogeneous media which include critical angle events. part III: appications in modeling, migration and inversion: Geophysical Prospecting, **34**, 180–207.

Wapenaar, C. P. A., van Gelovan, W., and van der Leij, T., 1987, Full prestack versus shot record migration: 69th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 761-764.

Wapenaar, C. P. A., Van Wijingaarden, A. J., van Gelovan, W., and van der Leij, T., 1999, Apparent AVA effects of fine layering: Geophysics, **64**, 1939–1948.

Weglein, A. B., and Stolt, R. H., 1999, Migration-inversion revisited(1999): The Leading Edge, **18**, no. 8, 950–952.

Wiggins, R., 1978, Minimum entropy deconvolution: Geoexploration, **16**, 21–35.

Wong, S., Foy, C., Gunter, W. D., and Jack, T., 1999, Injection of CO2 for enhanced energy recovery: Elsevier Sciences, New York.

Xie, X., and Wu, R., 1998, Improve the wide angle accuracy of screen method under large contrast: SEG Expanded Abstracts, pages 1811–1814.

Xu, S., Chauris, H., Lambare, G., and Noble, M. S., 1998, Common angle image gather: a new strategy for imaging complex media: 68 th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1538-1541.

Yilmaz, O., and Taner, M. T., 1994, Discrete plane-wave decomposition by least-mean-square-error method: Geophysics, **59**, 973–982.

Youzwishen, C. F., 2001, Non-linear sparse and blocky constrants for seismic inverse problems: MSc. thesis, the University of Alberta.

Zhou, B., M., M. I., and Greenhalgh, S. A., 1996, An accurate formulation of log-stretch dip moveout in the frequency-wavenumber domain: Geophysics, **61**, 815–820.

Appendix A

Propagation matrix for two-way wave equation

The two-way wave equation can be expressed in the wavenumber-frequency domain as:

$$\frac{\partial \tilde{\mathbf{Q}}}{\partial z} = \tilde{\mathbf{A}}\tilde{\mathbf{Q}}, \tag{A.1}$$

where

$$\begin{aligned}
\tilde{\mathbf{A}} &= \begin{pmatrix} 0 & i\omega\rho \\ -\tilde{H}_2/(i\omega\rho) & 0 \end{pmatrix}, && \text{(A.2)} \\
\tilde{\mathbf{Q}} &= \begin{pmatrix} -\tilde{P} \\ \tilde{V}_z \end{pmatrix}, && \text{(A.3)} \\
\tilde{H}_2 &= k_z^2 = k^2 - k_x^2 - k_y^2. && \text{(A.4)}
\end{aligned}$$

The eigenvalue composition of $\tilde{\mathbf{A}}$ reads,

$$\tilde{\mathbf{A}} = \tilde{\mathbf{L}}\tilde{\Lambda}\tilde{\mathbf{L}}^{-1}, \tag{A.5}$$

where

$$\begin{aligned}
\tilde{\mathbf{L}} &= \begin{pmatrix} 1 & 1 \\ \frac{-1}{\omega\rho H_1} & \frac{1}{\omega\rho H_1} \end{pmatrix}, && \text{(A.6)} \\
\tilde{\Lambda} &= \begin{pmatrix} -iH_1 & 0 \\ 0 & iH_1 \end{pmatrix}, && \text{(A.7)} \\
\tilde{L}^{-1} &= \frac{1}{2}\begin{pmatrix} 1 & -\omega H_1^{-1}\rho \\ 1 & \omega H_1^{-1}\rho \end{pmatrix}. && \text{(A.8)}
\end{aligned}$$

If the medium is homogeneous, the solution of equation A.1 is given by:

$$\tilde{\mathbf{Q}}(z) = \tilde{\mathbf{W}}(z, z_0)\tilde{\mathbf{Q}}(z_0), \tag{A.9}$$

where

$$\tilde{\mathbf{W}}(z, z_0) = e^{(\tilde{\mathbf{A}}\Delta z)}, \tag{A.10}$$

The phase-shift operator $\tilde{\mathbf{W}}(z, z_0)$ can be expanded to Taylor's:

$$\tilde{\mathbf{W}}(z, z_0) = \mathbf{I} + (\tilde{\mathbf{A}}\Delta z) + \frac{1}{2}(\tilde{\mathbf{A}}\Delta z)^2 + \cdots \tag{A.11}$$

Using the eigenvalue decomposition of $\tilde{\mathbf{A}}$, we can rewrite the series as

$$\begin{aligned}
\tilde{\mathbf{W}}(z, z_0) &= \mathbf{I} + \tilde{\mathbf{L}}(\tilde{\Lambda}\Delta z)\tilde{\mathbf{L}}^{-1} + \frac{1}{2}\tilde{\mathbf{L}}(\tilde{\Lambda}\Delta z)\tilde{\mathbf{L}}^{-1}\tilde{\mathbf{L}}(\tilde{\Lambda}\Delta z)\tilde{\mathbf{L}}^{-1} + \cdots \\
&= \tilde{\mathbf{L}}\Big[\mathbf{I} + (\tilde{\Lambda}\Delta z) + \frac{1}{2}(\tilde{\Lambda}\Delta z)^2 + \cdots\Big]\tilde{\mathbf{L}}^{-1} \\
&= \tilde{\mathbf{L}} \quad exp(\tilde{\Lambda}\Delta z) \quad \tilde{\mathbf{L}}^{-1} \\
&= \tilde{\mathbf{L}}\begin{bmatrix} exp(-i\tilde{H}_1\Delta z) & 0 \\ 0 & exp(i\tilde{H}_1\Delta z) \end{bmatrix}\tilde{\mathbf{L}}^{-1} \\
&= \begin{bmatrix} 1 & 1 \\ \frac{-1}{\omega\rho H_1} & \frac{1}{\omega\rho H_1} \end{bmatrix}\begin{bmatrix} exp(-i\tilde{H}_1\Delta z) & 0 \\ 0 & exp(i\tilde{H}_1\Delta z) \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1 & -\omega\tilde{H}_1^{-1}\rho \\ 1 & \omega\tilde{H}_1^{-1}\rho \end{bmatrix} \\
&= \begin{bmatrix} \tilde{W}_I(z, z_0) & \tilde{W}_{II}(z, z_0) \\ \tilde{W}_{III}(z, z_0) & \tilde{W}_{IV}(z, z_0) \end{bmatrix},
\end{aligned}$$

where

$$\begin{aligned}
\tilde{W}_I(z, z_0) &= \cos(\tilde{H}_1\Delta z), & \text{(A.12)} \\
\tilde{W}_{II}(z, z_0) &= \frac{i\omega\rho}{\tilde{H}_1}\sin(\tilde{H}_1\Delta z), & \text{(A.13)} \\
\tilde{W}_{III}(z, z_0) &= \frac{1}{(\omega\rho)^2}\tilde{H}_2\tilde{W}_{II}(z, z_0), & \text{(A.14)} \\
\tilde{W}_{IV}(z, z_0) &= \tilde{W}_I(z, z_0), & \text{(A.15)}
\end{aligned}$$

where

$$\tilde{H}_1 = \sqrt{\tilde{H}_2} \;=\; \sqrt{k^2 - k_x^2 - k_y^2} \tag{A.16}$$

The matrix $\tilde{\mathbf{W}}(z, z_0)$ is called the propagation matrix for the two-way wave equation. General form of the propagation matrix for arbitrary inhomogeneous media is given by Wapenaar and Berhout (Wapenaar and Berkhout, 1986a).

Appendix B

Relationship between pressure, acoustic impedance and particle velocity for plane waves

The stress-displacement equation for acoustic media can be written as

$$P = -K\nabla \cdot \mathbf{u}, \tag{B.1}$$

where P is the pressure, K is bulk modulus, and **u** is the displacement vector. The bulk modulus of a acoustic medium is calculated by

$$K = \rho c^2, \tag{B.2}$$

where ρ is the density and c is the wave speed.

Assuming a plane wave is propagating along z direction, we can simplify the stress-displacement equation B.1 into:

$$\begin{aligned} P &= -\rho c^2 \cdot u_z \\ &= -\rho c(c \cdot \frac{\partial u}{\partial z}) \\ &= -\rho c(\frac{\partial z}{\partial t} \cdot \frac{\partial u}{\partial z}) \\ &= -\rho c \cdot \frac{\partial u}{\partial t} \\ &= -\rho c \cdot v \\ &= -I \cdot v, \end{aligned} \tag{B.3}$$

where I is the acoustic impedance, and v is the particle velocity.

It is clear that the above equation can also be used for a plane wave in an arbitrary direction by just rotating the axis.

Appendix C

Imaging Jacobian for common-azimuth migration

This Appendix gives the imaging Jacobian for common-azimuth migration without considering the amplitude correction term of the vertical wavenumber for the common-azimuth migration. For the formulas including the amplitude correction, please refer to Sava and Biondi (2001b).

The vertical wavenumber of common-azimuth downward continuation can be expressed as the sum of two square roots:

$$k_z = SSR_1 + SSR_1, \tag{C.1}$$

where

$$\begin{aligned} SSR_1 &= \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}\left[(k_{m_x} - k_{h_x})^2 + (k_{m_y} - \hat{k}_{h_y})^2\right]}, & \text{(C.2)} \\ SSR_2 &= \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}\left[(k_{m_x} + k_{h_x})^2 + (k_{m_y} + \hat{k}_{h_y})^2\right]}, & \text{(C.3)} \end{aligned}$$

where $\hat{k}_{h_y}$ is the stationary phase point of the crossline offset wavenumber:

$$\begin{aligned} \hat{k}_{h_y} &= k_{m_y} \frac{\sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} + k_{h_x})^2} - \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} - k_{h_x})^2}}{\sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} + k_{h_x})^2} + \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} - k_{h_x})^2}} \\ &= k_{m_y} \frac{m-n}{m+n}, & \text{(C.4)} \end{aligned}$$

where

$$\begin{aligned} m &= \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} + k_{h_x})^2}, \\ n &= \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} - k_{h_x})^2} \end{aligned}$$

The inverse of imaging Jacobian is

$$\begin{aligned} (\frac{d\omega}{dk_z})^{-1} &= \frac{dk_z}{d\omega}|_{k_{h_x}} \\ &= \frac{1}{2}\frac{\frac{2\omega}{v^2} - \frac{1}{2}(k_{m_y} - \hat{k}_{h_y}) \cdot (-\frac{d\hat{k}_{h_y}}{d\omega})}{SSR_1} + \frac{1}{2}\frac{\frac{2\omega}{v^2} - \frac{1}{2}(k_{m_y} + \hat{k}_{h_y}) \cdot (\frac{d\hat{k}_{h_y}}{d\omega})}{SSR_2} \\ &= \frac{\frac{\omega}{v^2} + \frac{1}{4}(k_{m_y} - \hat{k}_{h_y}) \cdot \frac{d_{\hat{k}_{h_y}}}{d\omega}}{SSR_1} + \frac{\frac{\omega}{v^2} - \frac{1}{4}(k_{m_y} + \hat{k}_{h_y}) \cdot \frac{d_{\hat{k}_{h_y}}}{d\omega}}{SSR_2} \end{aligned} \qquad (C.5)$$

The derivative of the stationary wavenumber over frequency is calculated as follows:

$$\frac{d\hat{k}_{h_y}}{d\omega} = k_{m_y}\frac{(m-n)'(m+n) - (m-n)(m+n)'}{(m+n)^2}, \qquad (C.6)$$

where

$$\begin{aligned} (m-n)' &= \left(\sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} + k_{h_x})^2} - \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} - k_{h_x})^2}\right)' \\ &= \frac{\omega}{v^2} \cdot \left(\frac{1}{\sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} + k_{h_x})^2}} - \frac{1}{\sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} - k_{h_x})^2}}\right) \\ &= \frac{\omega}{v^2} \cdot \frac{n-m}{mn}, \end{aligned} \qquad (C.7)$$

and similarly,

$$(m+n)' = \frac{\omega}{v^2} \cdot \frac{n+m}{mn}. \qquad (C.8)$$

Using equation C.7 and C.8 in equation C.6 leads to

$$\frac{d\hat{k}_{h_y}}{d\omega} = k_{m_y} \cdot \frac{\omega}{v^2} \cdot \frac{2}{mn} \cdot \frac{n-m}{m+n}. \qquad (C.9)$$

Inserting equation C.9 and C.4 into equation C.5 yields

$$J_{k_{h_x}}^{-1} = \frac{dk_z}{d\omega} = \frac{\omega}{v^2}\left\{\frac{1 + k_{m_y}^2 \cdot \frac{n-m}{m} \cdot \frac{1}{(m+n)^2}}{SSR_1} + \frac{1 - k_{m_y}^2 \cdot \frac{n-m}{n} \cdot \frac{1}{(m+n)^2}}{SSR_2}\right\}, \qquad (C.10)$$

which provides the basis for true amplitude common-azimuth migration. With equation C.10, the derivative of vertical wavenumber over temporal frequency can be calculated and used to correct the downward continued wavefield before the imaging condition is applied.

For a flat reflector, the inverse imaging Jacobian can be simplified by using the fact that $SSR_1 = SSR_2 = \omega \cos\theta / v$, and $k_{m_y} = 0$. Then the inverse of the imaging Jacobian is

$$J_{k_{h_x}}^{-1} = \frac{2}{v \cos\theta} \tag{C.11}$$

The inverse imaging Jacobian can also be derived in terms of the inline offset ray parameter p_{h_x}. Some algebra leads to

$$\begin{aligned} J_{p_{h_x}}^{-1} = \frac{1}{SSR_1}\Big\{\frac{\omega}{v^2} + \frac{1}{4}\Big[(k_{m_x} - \omega p_{h_x})p_{h_x} + k_{m_y}^2 \cdot \frac{4}{M} \cdot \frac{1}{(M+N)^2}\Big(\big(\frac{\omega^2}{v^2} - \\ \frac{1}{4}\omega p_{h_x}^2\big)(N-M) - \frac{1}{4}k_{m_x}p_{h_x} \cdot \frac{M^2+N^2}{M+N}\Big)\Big]\Big\} \\ + \frac{1}{SSR_2}\Big\{\frac{\omega}{v^2} - \frac{1}{4}\Big[(k_{m_x} + \omega p_{h_x})p_{h_x} + k_{m_y}^2 \cdot \frac{4}{N} \cdot \frac{1}{(M+N)^2}\Big(\big(\frac{\omega^2}{v^2} - \\ \frac{1}{4}\omega p_{h_x}^2\big)(N-M) - \frac{1}{4}k_{m_x}p_{h_x} \cdot \frac{M^2+N^2}{M+N}\Big)\Big]\Big\}, \end{aligned} \tag{C.12}$$

where

$$\begin{aligned} M &= \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} + \omega \cdot p_{h_x})^2}, \\ N &= \sqrt{\frac{\omega^2}{v^2} - \frac{1}{4}(k_{m_x} - \omega \cdot p_{h_x})^2} \end{aligned}$$

For a flat reflector, the inverse imaging Jacobian can be simplified into

$$J_{p_{h_x}}^{-1} = \frac{2\cos\theta}{v}. \tag{C.13}$$

It is easy to verify that for flat reflectors

$$J_{p_{h_x}}^{-1} \cdot J_{k_{h_x}}^{-1} = \frac{4}{v^2} \tag{C.14}$$

Equation C.13 tells us that the amplitude of a flat reflector can be corrected by the cosine of the incident angle after migration.